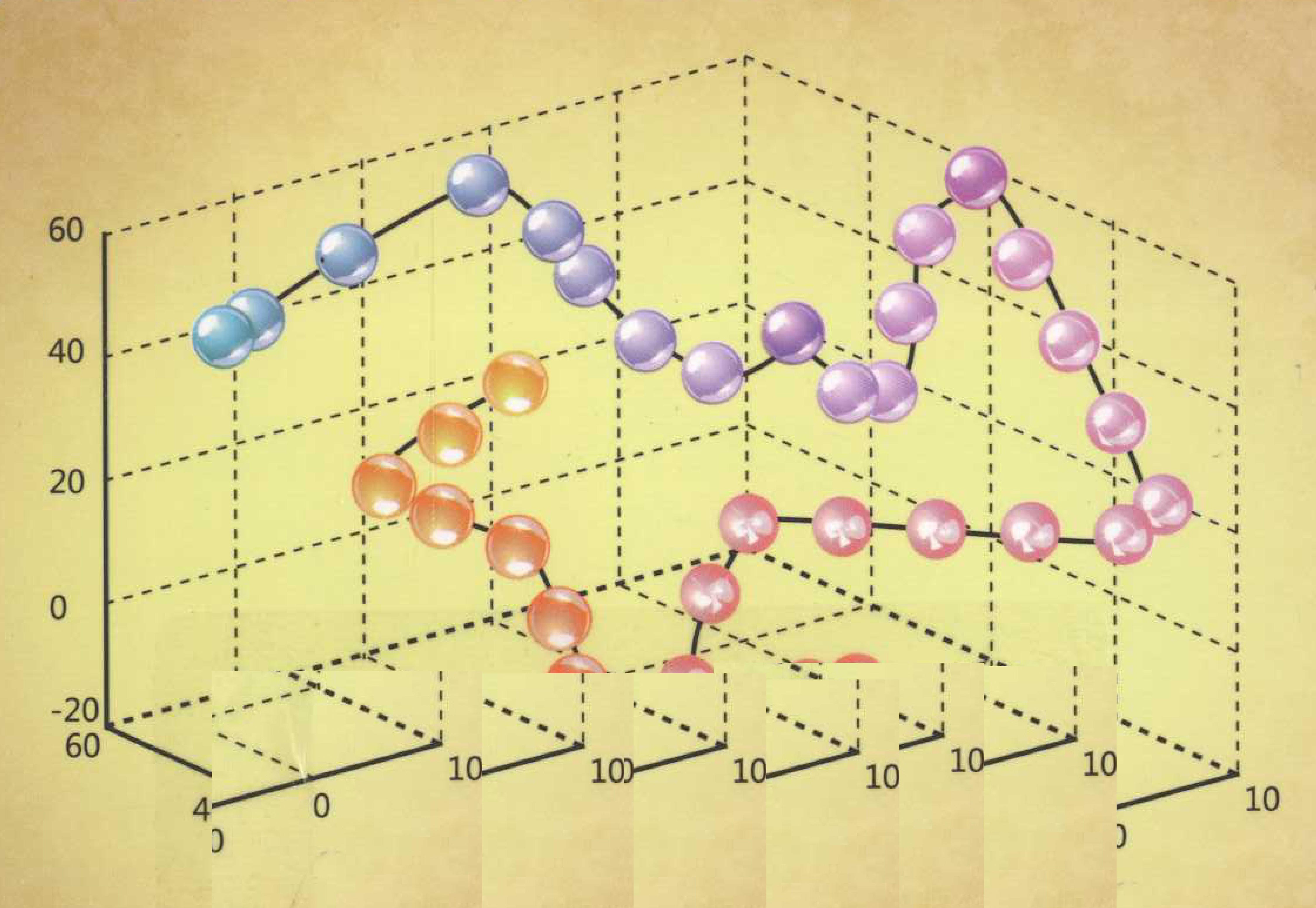
图灵数学·统计学丛书 44





随机过程

[日] 伊藤 清 著
刘璋温 译
郑绍濂 审校



随机过程

本书是日本著名数学家伊藤清的著作，是随机过程方面的经典名著，篇幅短小，叙述精辟，具有较高的理论水平。书中以简练的笔法介绍了随机过程论的主要方面，包括可加过程、平稳过程和Markoff过程，并概述了一维扩散过程。具有初步概率论和泛函分析知识的读者，可以借此快速掌握随机过程的基本理论。



伊藤 清(1915－2008)　日本数学家，日本学士院院士，世界级概率论大师。他因在概率论方面的奠基性工作而获1987年的沃尔夫奖，并于1998年获得京都基础科学奖，2006年获得首届高斯奖。伊藤清的工作集中于概率论，特别是随机分析领域，他被誉为"现代随机分析之父"，因他命名的理论有伊藤过程、伊藤公式和伊藤微积分。他的研究对其他学科尤其是金融数学产生了深远影响。

延伸阅读

Sheldon M. Ross 《应用随机过程：概率模型导论(第9版)》 中文版及英文影印版

Samuel Karlin, Howard M. Taylor 《随机过程初级教程(第2版)》 中文版及英文影印版



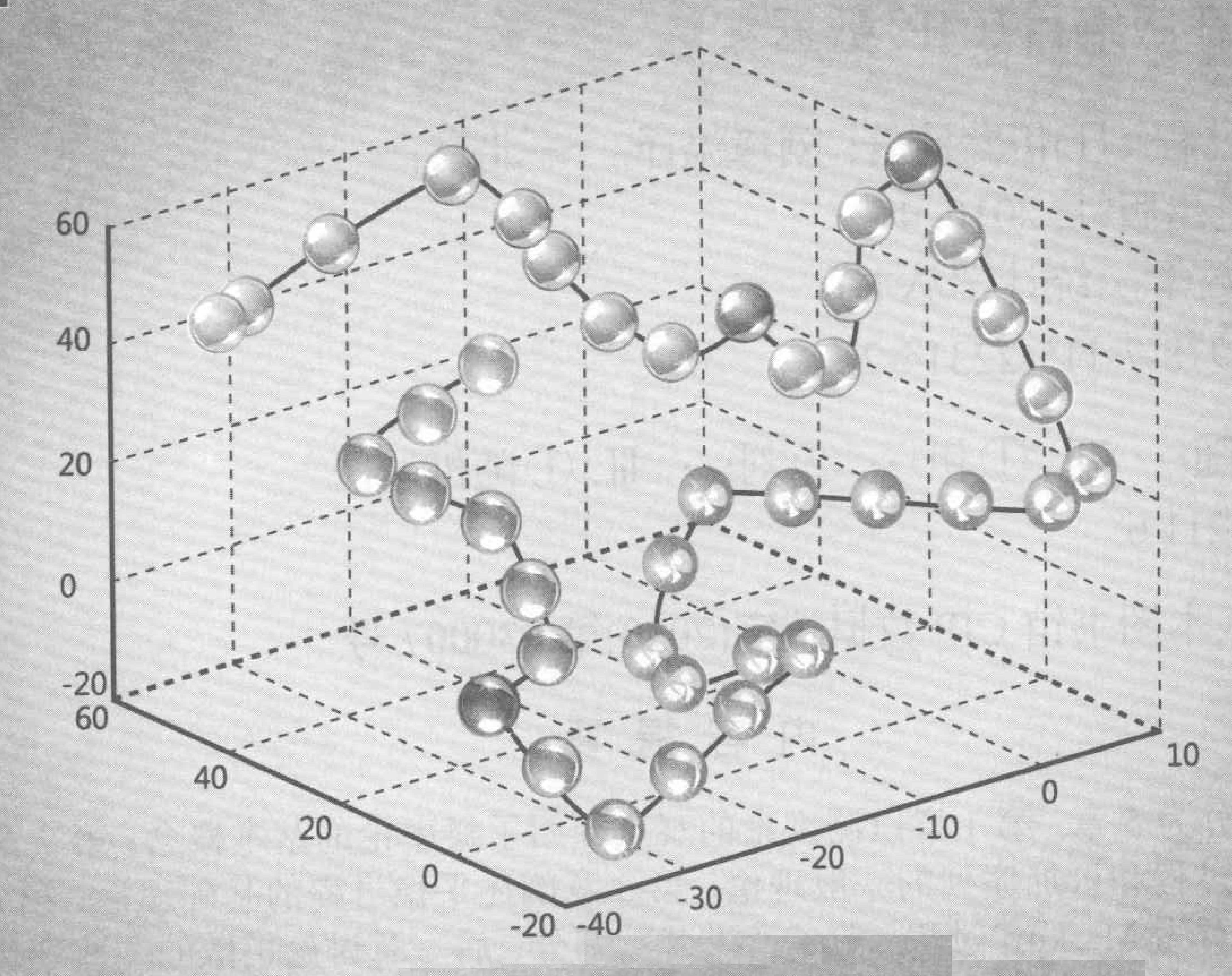
本书相关信息请访问：图灵网站 http://www.turingbook.com
读者/作者热线：(010)51095186
反馈/投稿/推荐信箱：contact@turingbook.com

分类建议：数学/概率
人民邮电出版社网址 www.ptpress.com.cn



ISBN 978-7-115-22314-2
定价:25.00元

TURING 图灵数学·统计学丛书 44



随机过程

[日] 伊藤清 著

刘璋温 译

郑绍濂 审校

人民邮电出版社

北京

图书在版编目(CIP)数据

随机过程/(日)伊藤清著; 刘璋温译. —北京:
人民邮电出版社, 2010. 4
(图灵数学・统计学丛书)
ISBN 978-7-115-22314-2

Ⅰ.①随… Ⅱ.①伊… ②刘… Ⅲ.①随机过程
Ⅳ.①O211.6

中国版本图书馆 CIP 数据核字(2010)第 030007 号

内 容 提 要

本书分为 5 章. 第 1 章以测度论的观点介绍了概率论的基本概念; 第 2 章叙述可加过程和可加序列的一般理论; 第 3 章阐述平稳过程的基础理论; 第 4 章和第 5 章为 Markoff 过程, 前一章讲基础部分, 后一章讨论扩散的一些现代理论和方法.

本书可供高等院校数学系、物理系等相关专业师生及工程师作参考.

图灵数学・统计学丛书

随机过程

◆ 著　[日] 伊藤 清
译　刘璋温
审校　郑绍濂
责任编辑　明永玲

◆ 人民邮电出版社出版发行　北京市崇文区夕照寺街 14 号
邮编 100061　电子函件 315@ptpress.com.cn
网址: http://www.ptpress.com.cn
北京隆昌伟业印刷有限公司印刷

◆ 开本: 880×1230 1/32
印张: 6.5
字数: 208 千字　2010 年 4 月第 1 版
印数: 1–3 000 册　2010 年 4 月北京第 1 次印刷

著作权合同登记号 图字: 01-2008-3847 号

ISBN 978-7-115-22314-2

定价: 25.00 元

读者服务热线: (010) 51095186　印装质量热线: (010) 67129223

反盗版热线: (010) 67171154

版 权 声 明

译 者 序

本书是根据伊藤清所著的《随机过程 I》和《随机过程 II》翻译的. 伊藤的原著是日本岩波书店从 1957 年起连续出版的一套"岩波讲座——现代应用数学"基础篇中的两本.

随机过程是概率论的一个主要组成部分, 它的研究与现代科学技术有密切的关系. 本书以不大的篇幅用简练的笔法对随机过程论的几个主要方面和一些最新成就作了精练和严格的阐述. 具有初步概率论和泛函分析知识的读者, 通过阅读这本书, 可以较快地掌握现代随机过程的基本理论和发展方向.

本书的翻译工作起于 E. B. Dynkin 教授来我国讲学, 他通过译者的介绍了解到原书的内容后, 非常称赞, 并建议早日把它翻译出来. 在他的指导下,《随机过程 I》的俄译本已于 1960 年出版. 本译稿是在 1958 年 6 月底完成的, 其中第 4 章的前几节为余潜修先生所译. 随后译稿一直流传于北京大学、复旦大学和南开大学, 作为专门化的参考教材. 这次承复旦大学郑绍濂同志和南开大学王梓坤同志分别整理前三章和后两章, 并由郑绍濂同志统一校订. 校样排出后, 译者重新校阅了一遍, 根据原文又作了一些修正.

本书的出版与上列许多同志的帮助分不开, 译者在此对他们表示深切的感谢.

刘璋温

1961 年 4 月北京

前　　言

1957 年, 岩波书店曾将拙著《随机过程 I》和《随机过程 II》作为 “岩波讲座 —— 现代应用数学” 中的两个分册出版, 现又作为单行本再次发行. 本书概括介绍了随机过程的三部分重要内容, 即可加过程、平稳过程和 Markoff 过程, 还讲解了一维扩散过程. 特别是针对一维扩散过程, 本书第一次以随机分析的方式规范介绍了有关局部构造和边界点分类的内容, 而在当时这些才刚刚由 William Feller 发现. 为此, 我本人感到颇为自豪. 然而, 从那时起到现在本书再版, 已经过去半个世纪, 身为著者确实感慨颇深.

本书首版前, 拙著《概率论》(岩波书店, 1953) 中也概括讲解了有关可加过程、平稳过程和 Markoff 过程的内容. 本书首版后, 在以下两本英文拙著中还详细讲解了可加过程和 Markoff 过程:

Lectures on Stochastic Processes, Tata Institute of Fundamental Research, Bombay, 1960

Stochastic Processes, edited by Ole E. Barndorff-Nielsen and Ken-iti Sato, Springer, 2004 (原本作为 Aarhus 大学讲义于 1969 年出版)

在以上列出的 Tata 研究所的讲义录以及我与 H. P. McKean 合著的另外两本书

Diffusion Processes and their Sample Paths, Springer, 1965;

Classics in Mathematics, Springer, 1996

的合订本中还详细讲解了一维扩散过程的局部构造. 在该合订本的 4.6 节中, 用概率论的方法证明了一维扩散过程中齐次方程解的边界附近的具体性质. 为便于不太熟悉概率论的读者理解, 本书的 §60 和 §61 用分析的方法给出其证明.

在编写本书时, 随机过程的规范研究才刚刚兴起. 正如 “后记” 所述, 在当时相关文献信息极为不足. 在随后的半个世纪中, 随机过程研究得到极大发展, 国内外都出版了大量优秀的相关书籍. *Stochastic Processes*,

Springer, 2004 一书的著者前言和编者序中都列举了一些相关书目.

本书首次出版后, 在 E. B. Dynkin 的指导下, A. D. Wentzell 将本书翻译成俄文, 并分别于 1960 年和 1963 年在莫斯科出版了第 I 册和第 II 册. 此外, 在 1959 年, 已故耶鲁大学教授角谷静夫注意到本书记述的一维扩散过程, 建议当时还是耶鲁大学研究生院学生的伊藤雄二把第 II 册翻译为英文. 最后, 油印了打字机原稿, 并分发给耶鲁大学校内外的数学研究人员. 历经半个世纪, 本书的两个分册都由伊藤雄二翻译为英文, 并作为美国数学学会数学翻译丛书中的 1 卷在*Essentials of Stochastic Processes*系列下出版. 此事令我极为欣喜. 伊藤雄二辛勤翻译期间, 福岛正俊和渡边信三仔细订正了首版中的错字等细小错误, 还替我适当修订了第 4 章和第 5 章的部分内容. 此次再版采用了与英译时相同的修订. 这里再次对福岛和渡边两位的努力以及岩波书店吉田宇一的帮助表示衷心感谢.

今年 6 月, 正在为本书新出版的单行本做准备时, 我终生的恩师弥永昌吉先生突然辞世. 而前述英译单行本的校对即将完成之时, 于 2004 年 8 月收悉角谷静夫先生的讣讯. 从 1942 年我开始写本领域论文以来, 两位先生就经常热情地给予我建议和鼓励, 到现在已历经 60 多年. 当然, 本书首版时我也得到了两位先生的极大帮助. 两位先生不能喜见装订精美的英文版和日文版新书, 甚为遗憾. 追忆两位先生生前的身影, 我衷心地为他们祈求冥福.

著　者

于 2006 年 12 月

目　录

第1章　基本概念

§1　测度论观点下的概率论 (1) 直观的背景

假设甲、乙二人掷硬币, 以先掷出正面者为胜. 现在试让甲先掷并考虑下面两个问题.

(i) 甲胜的概率是多少?

(ii) 直到决定胜负为止所掷的平均次数是多少?

首先, 观察这种比赛, 它可能出现哪些结果. 若以 O 表示掷出正面, U 表示掷出反面, 则可能产生的结果为

$$\left\{\begin{array}{l}\omega_1 = \mathrm{O} \\ \omega_2 = \mathrm{UO} \\ \omega_3 = \mathrm{UUO} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \omega_n = \underbrace{\mathrm{UU}\cdots\mathrm{U}}_{(n-1)}\mathrm{O} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \omega_\infty = \mathrm{UUU}\cdots\end{array}\right\} \tag{1.1}$$

ω_1 是甲先掷出正面而结束比赛的情形, ω_2 是甲掷出反面, 乙掷出正面而结束比赛的情形. ω_n 是直到第 $n-1$ 次为止二人都掷出反面, 在第 n 次才掷出正面而结束比赛的情形, 因此可由 n 的奇偶来决定是甲或是乙获胜. 最后, ω_∞ 是甲、乙双方总是掷出反面的情形, 这时比赛将无限继续下去. 由于 $\omega_1,\omega_2,\cdots,\omega_\infty$ 是这种比赛经过的各种样本, 所以叫做**样本点**(sample point), 其全体的集合 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_\infty\}$ 就叫做**样本空间**(sample space).

现在来考虑各个样本点的概率. 第 1 次掷出正面或者反面, 这都是同等可能的, 所以 ω_1 的概率是 1/2, 而剩下的 $\{\omega_2,\omega_3,\cdots,\omega_\infty\}$全体的概率就应该是 1/2. 又由同样的理由, 后者的概率 1/2 应平均地分给 ω_2 和剩下

的 $\{\omega_3, \omega_4, \cdots, \omega_\infty\}$, 即各为 1/4. 据此可知, 分布于各个点 $\omega_1, \omega_2, \cdots, \omega_\infty$ 上的概率分别为 1/2, 1/4, 1/8, $\cdots$,0.

以 $P(\omega)$ 表示分布于 ω 上的概率. 于是就有

$$\begin{aligned} &P(\omega_1) = 1/2, P(\omega_2) = 1/4, \cdots, \\ &P(\omega_n) = 1/2^n, \cdots, P(\omega_\infty) = 0. \end{aligned} \tag{1.2}$$

又若 E 为 Ω 的子集合, 则分布于 E 上的概率是分布于 E 的各点上的概率的总和, 即

$$P(E) = \sum_{\omega \in E} P(\omega). \tag{1.3}$$

这样一来, 就得到了一个集合函数 $P(E)$. 这个集合函数叫做**概率分布**(probability distribution).

其次, 考虑 (ii) 的问题. 直到决定胜负为止的次数可由各个样本点唯一决定. 比如 ω_1 时为 1, ω_2 时为 2, ω_n 时为 n. 这就是定义在样本空间上的函数, 以 $x(\omega)$ 来表示. 在样本空间上如此定义的函数叫做**随机变量**(random variable). 问题 (ii) 就是求随机变量 $x(\omega)$ 的平均值. 平均值有各种各样的定义方法, 而最普遍被采用的是以下的所谓**数学期望**(expectation):

$$E(x) = \sum_{\omega \in \Omega} x(\omega)P(\omega) = \sum_{n=1}^{\infty} nP(\omega_n) = \sum_{n=1}^{\infty} n \cdot \frac{1}{2^n} = 2. \tag{1.4}$$

现在转到问题 (i) 上. 所谓甲胜就是上述的 $x(\omega)$ 为奇数的情形, 因此可用条件 "$x(\omega)=$ 奇数" 来表示. 这种可以由有关样本点 ω 的条件来表示的事情就叫做**事件**(event). 事件的概率规定为分布于满足该条件的样本点全体的集合 E [这叫做该事件的**外延**(extension)] 上的概率 $P(E)$. 因此

$$P(\text{甲胜}) = P(E) = \sum_{n=0}^{\infty} P(\omega_{2n+1}) = \sum_{n=0}^{\infty} \frac{1}{2^{2n+1}} = \frac{2}{3}. \tag{1.5}$$

现在把上例中的本质部分抽出来看一看. 首先, 作为基础的东西就是命名为样本空间的集合 Ω 和它上面的概率分布 P. 此外, 还有一个定义在 Ω 上而称为随机变量的函数 ①, 它的数学期望由 (1.4) 左端的等式来定

① 随机变量的严格的定义将在下面讲到. 此处的定义只适用于样本空间只包含可数个元素的情形. —— 校者注

义. 与 Ω 里的点有关的条件叫做事件, 而该事件的概率就可规定为 P 对其外延 E 的值 $P(E)$. 但在这种一般的场合下, 如何来给出概率分布就需要讨论一下. 由上面的例子可知, 假如 Ω 为一可数集合, 只要给予各个样本点的概率使得其总和为 1, 则就可由 (1.3) 给出它的概率分布. 但是, 在把直线上的点集、平面上的点集以及更一般的集合 (例如做 Brown 运动的粒子的各种各样的运动途径) 分别看做是 Ω 的场合下, 按照上述朴素的方法就不可能给出概率分布. 然而, 若注意到概率分布的想法类似于质量分布, 而后者的数学理论就是测度论, 由此, 就很自然地可以设想, 测度论也将适用于概率分布.

正是基于这种出发点, 我们把概率论构造成为测度论的形式, 因而使得长期以来以常识或者直观为基础, 而且缺乏严格的逻辑推理的概率论真正成为数学理论的一个分支, 并且已在许多的应用上获得了有价值的成果.

§2 概 率 分 布

令 X 为一个集合, $\boldsymbol{B}$ 为由 X 的子集所构成的 Borel 集合体. 对于 $\boldsymbol{B}$ 的元素 (集合) 定义了 Lebesgue 测度 $P(E)$, 而且满足

$$P(X)=1 \tag{2.1}$$

时, 就称 P 为 $X(\boldsymbol{B})$ 上的**概率测度**(probability measuro) 或者**概率分布**, 或简称为**分布**.

首先, 作为最简单的情况, 试考虑 X 为有限集合时的情形. 令 X 的元素为 $x_1, x_2, \cdots, x_n$. 这时通常可将 $\boldsymbol{B}$ 取为由 X 所有的子集组成的集合 2^X. 若只包含一个点的集合的测度 $P(x_1), P(x_2), \cdots, P(x_n)$ 已给定, 则由测度的可加性, 对任意的 $E \subseteq X$ 的概率可以定义为

$$P(E)=\sum_{x\in E} P(x). \tag{2.2}$$

因此, 在这样的场合下, 只需给定点函数 $P(x)$ 就可决定概率分布了. 显然, $P(x)$ 应满足条件

$$P(x) \geqslant 0, \quad \sum P(x)=1. \tag{2.3}$$

特别是当 $P(x)$**不依赖于**x, 而且 $P(x) = 1/n$ 时, 就称 P 为**均匀分布**(uniform distribution).

其次, 对于 X 是可数无限集合的情形, 它与有限集合的情况大致相同. 但此时不再存在均匀分布了.

当 X 是实数集合 $\mathbb{R}^1$ 时, 问题就较困难了. 除了个别的特殊情况外, $\boldsymbol{B}$ 不可能取为由 $\mathbb{R}^1$ 的所有子集组成的集合 $2^{\mathbb{R}^1}$. $\boldsymbol{B}$ 的最自然的取法是取为包含所有开集的最小 **Borel 集合体**$\boldsymbol{B}^1$. 通常称 $\boldsymbol{B}^1$ 的元素为 **Borel 集合**. 令 P 为 $\mathbb{R}^1(\boldsymbol{B}^1)$ 上的概率分布. 对 $x \in \mathbb{R}^1$ 的任意邻域 U, 若它的 P 测度 $P(U)$ 为正时, 就称 x 为 P 的支撑点, 而这样的点的全体的集合就叫做 P 的**支集**(support). 特别若 $P(x) > 0$, 则显然 x 是 P 的支撑点, 对于这样的 x, 我们另外给它一个名称, 叫做 P 的**不连续点**(discontinuity point). P 的不连续点的全体 D 至多是一可数集合. 当 $P(D)$=1 时, 称 P 为**纯粹不连续**(purely discontinuous); 而 $P(D) = 0$ 时, 就称 P 为连续 (continuous). 作为比连续稍微强的条件, 有**绝对连续**(absolutely continuous) 的概念. 若 E 的通常的 Lebesgue 测度 $|E|$=0 时, 必有 $P(E) = 0$, 则称 P 为绝对连续. 这时 P 具有密度, 而且可写为

$$P(E) = \int_E f(x)\mathrm{d}x. \tag{2.4}$$

此处 $f(x)$ 应满足的条件是

$$f(x) \geqslant 0, \int_{\mathbb{R}^1} f(x)\mathrm{d}x = 1. \tag{2.5}$$

连续但不是绝对连续的概率分布叫做**奇异**(singular) 分布. 纯粹不连续分布、绝对连续分布和奇异分布是在 $\mathbb{R}^1(\boldsymbol{B}^1)$ 上的分布中重要的三种形状, 而任意的分布都可以由这三种形状的分布的凸组合 (convex combination) 来表示.(所谓 a 是 $a_1, a_2, \ldots, a_n$ 的凸组合, 就是指能够写成 $a = \sum c_i a_i (c_i \geqslant 0, \sum c_i = 1)$ 的形状.) 这就是 **Lebesgue 分解定理**.

例 1 **δ分布**(delta distribution) $\delta(\cdot; a)$. 这是纯粹不连续分布, 也就是上述的 D 只含有一点 a 的场合. 特别当 $a = 0$ 时就叫做**单位分布**(unitary distribution).

例 2 **二项分布**(binomial distribution) $B(\cdot; p, n), 0 < p < 1, n$ 是自然数. 这是纯粹不连续分布, 也就是由 $D = \{0, 1, 2, \cdots, n\}$,

$$P(k)=\binom{n}{k}p^k\cdot q^{n-k},\quad q=1-p,\quad k=0,1,2,\cdots,n \tag{2.6}$$

给定的分布. 因为 $P(k)$ 等于 $(p+q)^n$ 展开式中的第 k 项, 所以有二项分布的名称.

例 3　Poisson 分布(Poisson distribution) $P(\cdot;\lambda),\lambda>0$. 这是纯粹不连续分布, 也就是由 $D=\{0,1,2,\cdots\}$,

$$P(k)=\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!},\quad k=0,1,2,\cdots \tag{2.7}$$

给定的分布.

例 4　正态分布(normal distribution) $N(\cdot;a,v),a$ 是实数, $v>0$. 这是绝对连续分布, 密度由下式给定:

$$f(x)=\frac{1}{\sqrt{2\pi v}}\mathrm{e}^{-\frac{(x-a)^2}{2v}}. \tag{2.8}$$

例 5　Cauchy 分布(Cauchy distribution) $C(\cdot;a,c),a$ 是实数, $c>0$. 这是绝对连续分布, 密度由下式给定:

$$f(x)=\frac{c}{\pi}\cdot\frac{1}{c^2+(x-a)^2}. \tag{2.9}$$

当 X 为 m 维空间 $\mathbb{R}^m$ 时, 仍然可以照 $\mathbb{R}^1$ 的场合将结果推广. $\boldsymbol{B}$ 与一维的情况同样, 是包含所有开集的最小 Borel 集合体 $\boldsymbol{B}^m$. $\boldsymbol{B}^m$ 的元素叫做 Borel 集合. 与上述一样, 分布有三种形状, 并且 Lebesgue 分解定理成立.

例 6　δ 分布 $\delta(\cdot;\boldsymbol{a})$ 与一维的情况一样, 不过 $\boldsymbol{a}$ 是 $\mathbb{R}^m$ 的元素而已.

例 7　多项分布(multinomial distribution) $B(\cdot;\boldsymbol{p},n),n$ 是自然数, $\boldsymbol{p}=(p_1,p_2,\cdots,p_n),p_i\geqslant 0,\Sigma P_i=1$. 这是纯粹不连续分布, 而 D 是格点 $\boldsymbol{k}=(k_1,k_2,\cdots,k_m),\Sigma k_i=n,k_i\geqslant 0$ 的全体.

$$P(\boldsymbol{k})=\frac{n!}{k_1!k_2!\cdots k_m!}p_1^{k_1}p_2^{k_2}\cdots p_m^{k_m}. \tag{2.10}$$

例 8　正态分布 $N(\cdot;\boldsymbol{a},\boldsymbol{v})$, $\boldsymbol{a}$ 是 $\mathbb{R}^m$ 的元素, $\boldsymbol{V}$ 是对称正定 (狭义) 矩阵. 这是绝对连续的, 其密度为

$$f(\boldsymbol{x})=(2\pi)^{-\frac{m}{2}}(\det\boldsymbol{V})^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\boldsymbol{V}(\boldsymbol{x}-\boldsymbol{a}),(\boldsymbol{x}-\boldsymbol{a}))\right\}. \tag{2.11}$$

此处 $\boldsymbol{V}(\boldsymbol{x}-\boldsymbol{a})$ 是把线性变换 $\boldsymbol{V}$ 作用于向量 $\boldsymbol{x}-\boldsymbol{a}$ 而得的, (,) 表示内积.

如上所述, 从一维的 $\mathbb{R}^1$ 的情形可平行地推广到 m 维的 $\mathbb{R}^m$ 的情形, 但是从 m 维推广到无限维时却有本质上的困难. 比如在无限维空间上, 由于不存在像在 $\mathbb{R}^m$ 的场合下那样普通的 Lebesgue 测度, 所以不可能定义像绝对连续那样的概念. 若令 A 为任意的集合, 那么 $\mathbb{R}^A$ 就是 A 中的元素 α 所对应的实数 ξ 排成的 $\prod\limits_{\alpha}\xi_\alpha$ 全体的集合. 若 A 是有限集合, 则 $\mathbb{R}^A$ 是有限维空间; 但若 A 为无限集合, 则 $\mathbb{R}^A$ 是无限维的. 使 $\prod\limits_{\alpha}\xi_\alpha \to \alpha_0$ 的坐标 ξ_{α_0} 的映象叫做**射影**(projection), 以 p_{α_0} 来表示. $\prod\limits_{\alpha}\xi_a \to \mathbb{R}^A$ 中的点 $(\xi_{\alpha_1},\xi_{\alpha_2},\cdots,\xi_{\alpha_n})$($\alpha_i$ 都不相同) 的映象也叫做射影, 以 $p_{\alpha_1\cdots\alpha_n}$ 来表示. 若令 $E^{(n)}$ 为 n 维的 Borel 集合, 由形状如 $p^{-1}_{\alpha_1\cdots\alpha_n}(E^{(n)})$ 所表示的 $\mathbb{R}^A$ 的子集就叫做 $\mathbb{R}^A$ 的 **Borel 柱集**(cylinder set). 若以 $\boldsymbol{B}^A$ 表示包含所有 Borel 柱集的最小 Borel 集合体, 则 $\boldsymbol{B}^A$ 的元素就叫做 $\mathbb{R}^A$ 的 **Borel 集合**. 今令 P 为 $\mathbb{R}^A(\boldsymbol{B}^A)$ 上的分布. 对于不相同的 $\alpha_1,\alpha_2,\cdots,\alpha_n\in A$, 定义 $\mathbb{R}^n(\boldsymbol{B}^n)$ 上的分布 $P_{\alpha_1\cdots\alpha_n}$, 使得

$$P_{\alpha_1\cdots\alpha_n}(E^{(n)}) = P(p^{-1}_{\alpha_1\cdots\alpha_n}(E^{(n)})), E^{(n)}\in\boldsymbol{B}^n. \tag{2.12}$$

这就是分布 **P 的射影**. 考虑所有这类的 $P_{\alpha_1\cdots\alpha_n}$, 且令其全体为 $\mathfrak{P}$, 则 $\mathfrak{P}$ 满足下面所述的 Kolmogoroff **相容性条件**(consistency condition):

(K.1) 若令 $i(1),i(2),\cdots,i(n)$ 为 1,2, $\cdots,n$ 的排列, 则

$$P_{\alpha_{i(1)}\alpha_{i(2)}\cdots\alpha_{i(n)}}(E_{i(1)}\times E_{i(2)}\times\cdots\times E_{i(n)}) = P_{\alpha_1\cdots\alpha_n}(E_1\times E_2\times\cdots\times E_n).$$

(K.2) $P_{\alpha_1\alpha_2\cdots\alpha_n}(E^{(n-1)}\times\mathbb{R}^1) = P_{\alpha_1\alpha_2\cdots\alpha_{n-1}}(E^{(n-1)})$.

反之, 对于满足这两个条件的分布系 $\mathfrak{P}$, 存在唯一的 $\mathbb{R}^A(\boldsymbol{B}^A)$ 上的分布 P 满足 (2.12). 这就叫做 Kolmogoroff 定理 ①.

现在, 假设已对 A 的各个元素 α 定义了 $\mathbb{R}^1(\boldsymbol{B}^1)$ 上的分布 P_α. 这时若定义

$$P_{\alpha_1\cdots\alpha_n} = P_{\alpha_1}\times P_{\alpha_2}\times\cdots\times P_{\alpha_n}(\text{直积测度}), \tag{2.13}$$

则 $\mathfrak{P}=\{P_{\alpha_1\cdots\alpha_n}\}$ 满足上述的 (K.1) 和 (K.2). 因此, 按照 Kolmogoroff 定理, 可以由 $\mathfrak{P}$ 定出 $\mathbb{R}^A(B^A)$ 上的分布 P. 称它为 $P_\alpha,\alpha\in A_1$ 的**直积概率**

① 参见 A. H. Колмогоров著的中译本《概率论基本概念》, 丁寿田译, 商务印书馆, 1953. —— 校者注

测度(direct product probability measure), 且以 $\prod_{\alpha \in A} P_\alpha$ 来表示. 显然 P 是由

$$P(p_{\alpha_1}^{-1}(E_1) \cap p_{\alpha_2}^{-1}(E_2) \cap \cdots \cap p_{\alpha_n}^{-1}(E_n)) = \prod_{i=1}^{n} P_{\alpha_i}(E_i) \tag{2.14}$$

所表达的 $\mathbb{R}^A(\boldsymbol{B}^A)$ 上的分布. 同理, 当 P_α 为高维 (有限或者无限) 的分布时 (维数可以由 α 而变), 也可以定义其直积概率测度.

§3 测度论观点下的概率论 (2) 逻辑的构成

固定一个集合 Ω, 这集合叫做样本空间. 在 Ω 上取一 Borel 集合体 $\boldsymbol{B}$, 及 $\Omega(\boldsymbol{B})$ 上的概率分布 P, 则 $\Omega, \boldsymbol{B}, P$ 有一个总的名称, 叫做 Ω 上的**概率空间**(probability space), 并记以 $\Omega(\boldsymbol{B}, P)$.

因为 $\Omega(\boldsymbol{B}, P)$ 是一种测度空间, 所以在其上可以建立 Lebesgue 积分理论. $\Omega(\boldsymbol{B}, P)$ 上的可测实函数叫做**随机变量**(random variable). 令 $x(\omega)$ 为随机变量, 则称

$$\Phi(E) = P\{\omega / x(\omega) \in E\} \equiv P(x^{-1}(E)), \quad E \in \boldsymbol{B}^1 \tag{3.1}$$

为 x 的**分布**(distribution). 这就是 $\mathbb{R}^1(\boldsymbol{B}^1)$ 上的概率分布. 称

$$E(x) = \int_\Omega x(\omega) P(\mathrm{d}\omega) \tag{3.2}$$

为 $x(\omega)$ 的**期望**(expectation) 或**均值**(mean). 这可以利用 x 的分布写成形式 ①

$$E(x) = \int_{\mathbb{R}^1} \xi \Phi(\mathrm{d}\xi). \tag{3.3}$$

因为可测函数不一定是可积的, 所以随机变量的均值不一定存在.

将若干随机变量并列起来就可以定义随机向量. 令 A 为有限或者无限集合, 且对于 A 的每个元素 α 都有一随机变量 $x_\alpha(\omega)$ 与之对应. 此时, 若定义

① 参见 Paul R. Halmos 著的中译本《测度论》, 王建华译, 科学出版社, 1958, §39, 定理 3. —— 校者注

$$\boldsymbol{x}(\omega)=\prod_{\alpha\in A}x_\alpha(\omega), \tag{3.4}$$

则 $\boldsymbol{x}(\omega)$ 是定义于 Ω 上而在 $\mathbb{R}^A$ 中取值的函数. 并且在

$$E^{(A)}\in \boldsymbol{B}^A \Rightarrow \boldsymbol{x}^{-1}(E^{(A)})\in \boldsymbol{B} \tag{3.5}$$

的意义下为可测. $\boldsymbol{x}(\omega)$ 就叫做随机向量. 若 A 的势为 m, 则叫做 m 维随机向量. 特别是把二维随机向量 $(x_1(\omega),x_2(\omega))$ 写成 $x_1(\omega)+ix_2(\omega)$, 就得**复随机变量**. 对于随机向量来说, 其分布也可以由 (3.1) 来定义, 再对各个分量进行积分, 就可以定义它的均值向量.

令 φ 为由 $\mathbb{R}^A$ 到 $\mathbb{R}^B$ 内的映象, 并且满足

$$E^{(B)}\in \boldsymbol{B}^B \Rightarrow \varphi^{-1}(E^{(B)})\in \boldsymbol{B}^A. \tag{3.6}$$

这时称 φ 是 **Borel 可测的**(Borel measurable) 或称 **$\boldsymbol{B}$–可测的**(B-measurable), 或者更简单地称**可测**(measurable). 由随机向量的可测映象所产生的象是可测的. 也就是说, 若把上述的可测映象 φ 作用于取值于 $\mathbb{R}^A$ 内的随机向量 $x(\omega)$ 上, 就得到一个在 $\mathbb{R}^B$ 内取值的随机向量 $\varphi(\boldsymbol{x}(\omega))$. 这时 $\varphi(\boldsymbol{x}(\omega))$ 的均值向量可由

$$E[\varphi(\boldsymbol{x}(\omega))]=\int_{\mathbb{R}^A}\varphi(\xi)\Phi\ (\mathrm{d}\xi),\quad \Phi\text{是}\boldsymbol{x}(\omega)\text{的分布} \tag{3.7}$$

给定. 这样的 $\varphi(\boldsymbol{x}(\omega))$ 称为**关于 $\boldsymbol{x}(\omega)$ 是可测的**. 若有若干随机向量 $\boldsymbol{x}_\alpha(\omega),\alpha\in A$, 也可以把它们排列起来定义更高维的随机向量

$$\boldsymbol{x}(\omega)=\prod_{\alpha}\boldsymbol{x}_\alpha(\omega). \tag{3.8}$$

若 $\boldsymbol{x}(\omega)$ 的分布是 $\boldsymbol{x}_\alpha(\omega)(\alpha\in A)$ 的分布的直积分布, 则称 $\boldsymbol{x}_\alpha(\omega)(\alpha\in A)$ 为**独立**(independent). 这也可以由下述条件来表达.

对任意不相同的 $\alpha_1,\alpha_2,\cdots,\alpha_n\in A$, 都有

$$P\{\omega/\boldsymbol{x}_{\alpha_i}(\omega)\in E_i, i=1,2,\cdots,n\}=\prod_{i=1}^{n}P\{\omega/\boldsymbol{x}_{\alpha_i}(\omega)\in E_i\}. \tag{3.9}$$

设 $A=\sum\limits_{\lambda}A_\lambda$(直和), $\lambda\in\Lambda$, 且令 $x_\alpha(\omega)(\alpha\in A)$ 为独立, 则

$$\boldsymbol{y}_\lambda(\omega)=\prod_{\alpha\in A_\lambda}\boldsymbol{x}_\alpha(\omega),\quad \lambda\in\Lambda \tag{3.10}$$

也为独立. 又若令 $\boldsymbol{x}_\alpha(\omega)(\alpha \in A)$ 为独立, 且令 φ_α 为可测, 则 $\varphi_\alpha(\boldsymbol{x}_\alpha(\omega))$ $(\alpha \in A)$ 也为独立.

若 $x_1(\omega), x_2(\omega), \cdots, x_n(\omega)$ 是独立的复 (实) 随机变量, 则

$$E[x_1(\omega)\cdots x_n(\omega)] = E[x_1(\omega)]E[x_2(\omega)]\cdots E[x_n(\omega)]. \tag{3.11}$$

这叫做均值的可乘性.

关于随机变量 (或者向量) 的序列 $x_n(\omega)(n=1,2,\cdots)$ 收敛于 $x(\omega)$ 的定义, 有各种各样的方法. 一个最自然的方法是

$$P\{\omega/|x_n(\omega)-x(\omega)| \to 0\} = 1, \tag{3.12}$$

此处 $|\ |$ 表示向量的长度. 因为 $|x_n(\omega)-x(\omega)|$ 是 ω 的可测函数, 并且

$$\{\omega/|x_n(\omega)-x(\omega)| \to 0\} = \bigcap_p \bigcup_N \bigcap_{n>N} \left\{\omega/|x_n(\omega)-x(\omega)| < \frac{1}{p}\right\},$$

所以 (3.12) 左边的 $\{\omega/''\}$ 是可测集合. 而 (3.12) 即意味着它的 P–测度为 1. 这种收敛, 叫做 Ω 上的**几乎处处**(almost everywhere)**收敛**, 或概率为 1 的收敛, 或**几乎必然收敛**(almost sure convergence), 或**概收敛**, 并记为 $x_n \to x$(a.e.)①.

下面是较弱的一种收敛概念. 对所有的 $\varepsilon > 0$, 极限

$$P\{\omega/|x_n(\omega)-x(\omega)| > \varepsilon\} \to 0 \tag{3.13}$$

恒成立, 这叫做**依概率收敛**(convergence in probability), 并记为 $x_n \to x(P)$. 这时 x_n 的分布就以下一节中所解释的意义收敛于 x 的分布. 又若当 $E(|x_n - x|^p)$ 为有限时, 由

$$E(|x_n - x|^p) \to 0 \tag{3.14}$$

所定义的收敛, 叫做 ***p* 幂平均收敛**(mean convergence of the pth power). 当 $p=2$ 时, 称为**平均收敛**. 这种收敛的条件比依概率收敛的条件强.

§4 分布函数、特征函数、均值和方差

以 $\Phi, \Phi_1, \Phi_2, \cdots$ 表示 $\mathbb{R}^1(\boldsymbol{B}^1)$ 上的分布. 称函数

① a.e. 是 almost everywhere 的简写.

$$F(\xi)=\Phi(-\infty,\xi] \tag{4.1}$$

为 Φ 的**分布函数**(distribution function). $F(\xi)$ 应满足的条件是

(i) 非降: $\xi<\eta\Rightarrow F(\xi)\leqslant F(\eta)$,

(ii) 右连续: $F(\xi+0)=F(\xi)$,

(iii) $F(-\infty)=0,\qquad F(\infty)=1$.

反之, 若 $F(\xi)$ 满足以上条件, 则它是由

$$\Phi(E)=\int_E \mathrm{d}F(\xi)\quad \text{(Lebesgue-Stieltjes 积分)} \tag{4.2}$$

所决定的 $\mathbb{R}^1(\boldsymbol{B}^1)$ 上的分布 Φ 的分布函数. 现在若令 $x(\omega)$ 为随机变量, 则 $x(\omega)$ 的分布 Φ 的分布函数 $F(\xi)$ 可由等式

$$F(\xi)=P\{\omega/x(\omega)\leqslant\xi\} \tag{4.3}$$

给定. 这也叫做 $\boldsymbol{x(\omega)}$ **的分布函数**.

由以上可知, 分布 Φ 和分布函数 F 之间存在着一一对应关系, 所以可用分布函数来代表分布. 若以 $F,F_1,F_2,\cdots$ 表示对应于 $\Phi,\Phi_1,\Phi_2,\cdots$ 的分布函数, 则可将分布序列 $\{\Phi_n\}$ 的收敛定义如下.

若对于 $F(\xi)$ 的所有连续点 ξ, 均有 ①

$$F_n(\xi)\to F(\xi), \tag{4.4}$$

则规定为 $\Phi_n\to\Phi$. 这一条件能写成下述形式.

在 $\mathbb{R}^1$ 内稠密的点列 $\{\xi_m\}$ 上, 成立

$$F(\xi_m-0)\leqslant\varliminf_n F_n(\xi_m)\leqslant\varlimsup_n F_n(\xi_m)\leqslant F(\xi_m+0). \tag{4.4$'$}$$

或者, 对于任意的有界连续函数 $f(\xi)$, 成立极限关系

$$\int_{\mathbb{R}^1} f(\xi)\Phi_n(\mathrm{d}\xi)\to\int_{\mathbb{R}^1} f(\xi)\Phi(\mathrm{d}\xi). \tag{4.4$''$}$$

若

$$\inf_n \Phi_n[-a,a]\to 1\qquad (a\to\infty), \tag{4.5}$$

① 下面有关分布函数列收敛的一些结论, 其证明可参见 M. Loére: Probability Theory, Van Nostrand, 1955, §11. —— 校者注

则可以从 Φ_n 先取一子序列, 使之收敛于某一分布.

令

$$\varphi(z) = \int_{\mathbb{R}^1} e^{iz\xi} \Phi(d\xi), \quad -\infty < z < \infty, \tag{4.6}$$

即为 Φ 的 Fourier 变换, 此时称 φ 为 Φ 的**特征函数**(characteristic function). 若 Φ 是实随机变量 $x(\omega)$ 的分布, 则

$$\varphi(z) = E(e^{izx}). \tag{4.7}$$

这就叫做 $x(\omega)$ 的特征函数. 显然 $\varphi(z)$ 有下述的性质:

$$\varphi(0) = 1, \qquad |\varphi(z)| \leqslant 1, \tag{4.8}$$

$$\text{在} -\infty < z < \infty \text{内一致连续}. \tag{4.9}$$

非负定: 对于任意的复数 $a_1, a_2, \cdots, a_n$ 和任意的实数 $z_1, z_2, \cdots, z_n$,

$$\sum a_i \overline{a}_j \varphi(z_i - z_j) \geqslant 0. \tag{4.10}$$

反之, 一个非负定的函数 $\varphi(z)$, 在 $z = 0$ 连续且 $\varphi(0) = 1$, 则它必为某一分布 Φ 的特征函数. 这就是 **Bochner 定理**①.

令 $\varphi, \varphi_1, \varphi_2, \cdots$ 分别为 $\Phi, \Phi_1, \Phi_2, \cdots$ 的特征函数, 则成立下面的 Φ 和 φ 的对应关系.

(i) $\varphi_1(z) \equiv \varphi_2(z) \Leftrightarrow \Phi_1 = \Phi_2$.

(ii) $\varphi_n(z) \to \varphi(z)$ (对所有的 z) $\Leftrightarrow \Phi_n \to \Phi$.

(iii) 若 $\varphi_n(z)$ 在所有的 z 上收敛于某一函数 $\theta(z)$(并未假定 $\theta(z)$ 是某一分布的特征函数), 并且在 $z = 0$ 的某些邻域内一致收敛时, 则 $\theta(z)$ 就成为某一分布 Φ 的特征函数. 因此, 由 (ii) 得出 $\Phi_n \to \Phi$.

这些性质主要是由 P. Lévy 所获得的.

分布的**均值 (mean)** 和**方差 (variance)** 分别由

$$M(\Phi) = \int_{\mathbb{R}^1} \xi \Phi(d\xi), \quad V(\Phi) = \int_{\mathbb{R}^1} (\xi - M(\Phi))^2 \Phi(d\xi) \tag{4.11}$$

来定义. 可以这样说, 均值是分布的中心, 而方差表示分布的散布的程度. 若令 Φ 为 $x(\omega)$ 的分布, 则 $M(\Phi)$ 等于 $E(x)$, 而 $V(\Phi)$ 就等于 $E[(x -$

① 参见复旦大学数学系郑绍濂等编的《概率论与数理统计》, 上海科学技术出版社, 1961,§19. —— 校者注

$E(x))^2]$. 后者也记为 $V(x)$, 并叫做 x 的方差. 显然,

$$V(x) = E(x^2) - E(x)^2. \tag{4.12}$$

令 x_1, x_2 为独立随机变量, x 为其和. 令 x_1, x_2, x 的分布各为 Φ_1, Φ_2, Φ; 分布函数为 F_1, F_2, F; 特征函数为 $\varphi_1, \varphi_2, \varphi$; 方差为 V_1, V_2, V, 则有

$$\Phi(M) = \int_{\mathbb{R}^1} \Phi_1(M-\xi)\Phi_2(\mathrm{d}\xi) = \int_{\mathbb{R}^1} \Phi_2(M-\xi)\Phi_1(\mathrm{d}\xi), \tag{4.13}$$

此处
$$M - \xi = \{\eta - \xi/\eta \in E\},$$

$$F(\xi) = \int_{\mathbb{R}^1} F_1(\xi - \eta)\mathrm{d}F_2(\eta) = \int_{\mathbb{R}^1} F_2(\xi-\eta)\mathrm{d}F_1(\eta), \tag{4.14}$$

$$\varphi(z) = \varphi_1(z)\varphi_2(z) \qquad (\text{特征函数的可乘性}), \tag{4.15}$$

$$V = V_1 + V_2 \qquad (\text{方差的可加性}). \tag{4.16}$$

为了证明 (4.13), 令 c_M 为 M 的示性函数, 则因随机向量 (x_1, x_2) 的分布为 $\Phi_1 \times \Phi_2$, 所以

$$\begin{aligned}\Phi(M) &= E[c_M(x)] = E[c_M(x_1 + x_2)] \\ &= \int_{\mathbb{R}^1}\int_{\mathbb{R}^1} c_M(\xi_1 + \xi_2)\Phi_1(\mathrm{d}\xi_1)\Phi_2(\mathrm{d}\xi_2) \\ &= \int_{\mathbb{R}^1} \Phi_1(M - \xi_2)\Phi_2(\mathrm{d}\xi_2).\end{aligned}$$

其他的关系式也可仿此加以证明. 姑且脱离随机变量而来考虑任意的两个分布 Φ_1, Φ_2, 则由 (4.13) 所定义的 Φ 也是一个分布, 并以 $\Phi_1 * \Phi_2$ 来表示, 称它为 Φ_1 和 Φ_2 的卷积 (convolution). 若令 Φ_1, Φ_2, Φ 的特征函数为 $\varphi_1, \varphi_2, \varphi$, 则

$$\varphi(z) = \varphi_1(z)\varphi_2(z) \Leftrightarrow \Phi = \Phi_1 * \Phi_2. \tag{4.17}$$

若 Φ 为一分布, 定义 Φ 的负反 $\check{\Phi}$ 如下:

$$\check{\Phi}(M) = \Phi(-M), \quad -M = \{-\xi/\xi \in M\}. \tag{4.18}$$

$\check{\Phi}$ 的特征函数 $\check{\varphi}$ 即为

$$\check{\varphi}(z) = \overline{\varphi(z)}. \tag{4.19}$$

若 x 的分布为 Φ, 则 $-x$ 的分布即为 $\check{\Phi}$. 若 Φ 是 $\{\Phi_i\}$ 的凸组合, 则 φ 是同一组系数的 $\{\varphi_i\}$ 的凸组合. 由此可知, 特征函数全体的集合 C 具有下述的性质:

(i) C 中元素的凸组合仍属于 C,

(ii) $\varphi_1, \varphi_2 \in C \Rightarrow \varphi_1 \cdot \varphi_2 \in C$,

(iii) $\varphi \in C \Rightarrow \overline{\varphi} \in C$,

(iv) $\varphi \in C \Rightarrow |\varphi|^2 \in C$.

若令随机变量 x 的特征函数为 $\varphi(z)$, 则 $ax+b$ 的特征函数为 $\mathrm{e}^{ibz}\varphi(az)$.

现将在 §2 节所举的一维分布示例中分布的均值 M、方差 V 和特征函数 $\varphi(z)$ 列表给出, 见表 4-1.

表 4-1

Φ		M	V	$\varphi(z)$
δ 分布	$\delta(\cdot;a)$	a	0	e^{iaz}
二项分布	$B(\cdot;p,n)$	np	npq	$(p\mathrm{e}^{iz}+q)^n$
Poisson 分布	$P(\cdot;\lambda)$	λ	λ	$\exp\{\lambda(\mathrm{e}^{iz}-1)\}$
正态分布	$N(\cdot;a,v)$	a	v	$\exp\{iaz-\frac{v}{2}z^2\}$
Cauchy 分布	$C(\cdot;a,c)$	没有	没有	$\exp\{iaz-c\|z\|\}$

因为 $(p\mathrm{e}^{iz}+q)^n \cdot (p\mathrm{e}^{iz}+q)^{n'} = (p\mathrm{e}^{iz}+q)^{n+n'}$, 所以由 (4.17) 得

$$B(\cdot;p,n) * B(\cdot;p,n') = B(\cdot;p,n+n').$$

同理可得

$$\delta(\cdot;a) * \delta(\cdot;a') = \delta(\cdot;a+a'),$$

$$P(\cdot;\lambda) * P(\cdot;\lambda') = P(\cdot;\lambda+\lambda'),$$

$$N(\cdot;a,v) * N(\cdot;a',v') = N(\cdot;a+a',v+v'),$$

$$C(\cdot;a,c) * C(\cdot;a',c') = C(\cdot;a+a',c+c').$$

又因当 $v \to 0$ 时, $\exp\left\{iaz-\frac{v}{2}z^2\right\} \to \mathrm{e}^{iaz}$, 所以由式 (4.10) 和式 (4.11) 之间的 (ii) 得 $N(\cdot;a,v) \to \delta(\cdot;a)$. 在这种意义下, δ 分布可以看做正态分布的退化情形. 同理 δ 分布又可以看做 Cauchy 分布的退化情形.

上面所得到的结果仍然可以照样推广到 m 维的分布上去. 此时, 分布函数就由下式所给定:

$$F(\xi_1,\cdots,\xi_m)=\Phi((-\infty,\xi_1]\times(-\infty,\xi_2]\times\cdots\times(-\infty,\xi_m]). \tag{4.20}$$

分布序列收敛的定义也与一维时相仿. 特征函数可由

$$\varphi(\boldsymbol{z})=\int_{\mathbb{R}^m}\mathrm{e}^{i(\boldsymbol{z},\boldsymbol{\xi})}\Phi(\mathrm{d}\boldsymbol{\xi}),\quad(\boldsymbol{z},\boldsymbol{\xi})=\sum_{v=1}^{m}z_\nu\xi_\nu \tag{4.21}$$

来定义, 而其性质就完全与一维的情况相同. 均值和方差分别成为均值向量和方差矩阵. 其分量可由

$$M_i=\int_{\mathbb{R}^m}\xi_i\Phi(\mathrm{d}\boldsymbol{\xi}),\quad V_{ij}=\int_{\mathbb{R}^m}(\xi_i-M_i)(\xi_j-M_j)\Phi(\mathrm{d}\boldsymbol{\xi}) \tag{4.22}$$

给出. 若 Φ 是随机向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 的分布, 则上述的 F,φ,M 和 V 分别为

$$F(\xi_1,\cdots,\xi_m)=P\{\omega/x_1(\omega)\leqslant\xi_1,\cdots,x_m(\omega)\leqslant\xi_m\},$$
$$\varphi(\boldsymbol{z})=E(\mathrm{e}^{i(\boldsymbol{z},\boldsymbol{x})}),$$
$$M=E(\boldsymbol{x}),\quad V_{ij}=E\{(x_i-E(x_i))(x_j-E(x_j))\}.$$

若就 §2 所举的 m 维分布的例子来求 M,V 和 φ, 则可列为表 4-2.

表 4-2

Φ	M	V	φ
$\delta(\cdot;\boldsymbol{a})$	$\boldsymbol{a}$	0	$\mathrm{e}^{i(\boldsymbol{z},\boldsymbol{a})}$
$B(\cdot;\boldsymbol{p},n)$	$n\cdot\boldsymbol{p}$	$\begin{cases}V_{\mu\mu}=np_\mu(1-p_\mu)\\V_{\mu\nu}=-np_\mu p_\nu(\mu\neq\nu)\end{cases}$	$\left(\sum_\nu p_\nu\mathrm{e}^{iz_\nu}\right)^n$
$N(\cdot;\boldsymbol{a},\boldsymbol{V})$	$\boldsymbol{a}$	$\boldsymbol{V}$	$\exp\{i(\boldsymbol{a},\boldsymbol{z})-\frac{1}{2}(\boldsymbol{Vz},\boldsymbol{z})\}$

在定义正态分布的时候, 假定了 $\boldsymbol{V}$ 是狭义正定对称矩阵. 当 $\boldsymbol{V}$ 是广义正定对称时, 若令 $\boldsymbol{V}_n=\boldsymbol{V}+\boldsymbol{I}/n$($\boldsymbol{I}$ 是单位矩阵), 则 $\boldsymbol{V}_n$ 成为狭义正定对称, 因而可以定义 $N_n=N(\cdot;\boldsymbol{a},\boldsymbol{V}_n)$. 这个分布的特征函数 $\varphi_n(\boldsymbol{z})$ 可由下式给出:

$$\varphi_n(\boldsymbol{z})=\exp\Big\{i(\boldsymbol{a},\boldsymbol{z})-\frac{1}{2}(\boldsymbol{V}_n\boldsymbol{z},\boldsymbol{z})\Big\}=\varphi(\boldsymbol{z})\exp\Big\{-\frac{1}{2n}(\boldsymbol{z},\boldsymbol{z})\Big\},$$

此处

$$\varphi(\boldsymbol{z})=\exp\Big\{i(\boldsymbol{a},\boldsymbol{z})-\frac{1}{2}(\boldsymbol{Vz},\boldsymbol{z})\Big\}.$$

故 $\varphi_n(\boldsymbol{z})$ 在 $\boldsymbol{z}$ 的任意有界集合上一致收敛于 $\varphi(\boldsymbol{z})$, 因为 $\varphi_n(\boldsymbol{z})$ 是 N_n 的特征函数, 所以 $\varphi(\boldsymbol{z})$ 也就是某一分布 N 的特征函数, 而且 N 是分布序列 $\{N_n\}$ 的极限. 这个 N 记为正态分布 $N(\cdot;\boldsymbol{a},\boldsymbol{V})$. 当 $\boldsymbol{V}$ 特别是狭义正定时, 这就与上述的定义相一致. 若 $\boldsymbol{V}$ 不是狭义正定即 $\det\boldsymbol{V}=0$ 时, 那么这个正态分布就叫做**退化的正态分布**. 这时, 这个分布的支集不是整个空间 $\mathbb{R}^m$, 而是其中的某些超平面. 但即使是在退化的场合, $\boldsymbol{a}$ 和 $\boldsymbol{V}$ 仍然分别是 $N(\cdot;\boldsymbol{a},\boldsymbol{V})$ 的均值和方差.

最后, 就无限维空间 $\mathbb{R}^A(\boldsymbol{B}^A)$ 上的分布, 定义其均值向量 $\boldsymbol{M}$、方差矩阵 $\boldsymbol{V}$ 和特征函数 $\varphi(z)$. $\boldsymbol{M}$ 和 $\boldsymbol{V}$ 完全与 m 维的场合的定义相同. 特征函数与以前稍有不同的地方. 在 m 维的情况下, $\boldsymbol{z}$ 也是 m 维的任意向量, 但在无限维 ($\mathbb{R}^A$) 的情况下, $\boldsymbol{z}$ 就只取几乎所有的 (即除有限个例外的意思) 坐标为 0 的 $\mathbb{R}^A$ 的元素. 令这样的点的全体为 $\mathbb{R}_0^A$. 当 $\boldsymbol{z}\in\mathbb{R}_0^A,\boldsymbol{\xi}\in\mathbb{R}^A$ 时, 就可以定义 $(\boldsymbol{z},\boldsymbol{\xi})=\sum\limits_{a} z_\alpha\xi_\alpha$. 因为右边实际上是有限和, 所以不会发生收敛与否的问题. 对于 $\mathbb{R}^A(\boldsymbol{B}^A)$ 上的分布 P, 其特征函数 $\varphi(\boldsymbol{z})$ 就定义为

$$\varphi(\boldsymbol{z})=\int_{\mathbb{R}^A}\mathrm{e}^{i(z,\xi)}P(\mathrm{d}\xi),\quad z\in\mathbb{R}_0^A.$$

在 $\varphi(\boldsymbol{z})$ 中留下有限个坐标, 然后令剩下的都为 0, 就得 $\varphi(z)$ 的**截口**(section). P 的特征函数的截口就是 P 的射影 (2.12) 的特征函数. 据此, **Kolmogoroff 定理**可以改述如下.

若 $\boldsymbol{z}\in\mathbb{R}_0^A$ 的函数 $\varphi(\boldsymbol{z})$ 的任意截口是特征函数, 则 $\varphi(\boldsymbol{z})$ 是 $\mathbb{R}^A(\boldsymbol{B}^A)$ 上的某个分布的特征函数. 并且这分布是唯一决定的.

利用这点, 就可以定义**无限维的正态分布**. 令 $\boldsymbol{M}$ 为 $\mathbb{R}^A$ 的任意的元素, $\boldsymbol{V}$ 为 $\mathbb{R}^{A\times A}$ 的元素, 并且满足

$$V_{\alpha\beta}=V_{\beta\alpha}\quad (V_{\alpha\beta}\text{ 为 }\boldsymbol{V}\text{ 的坐标}),$$

$$(\boldsymbol{V}\boldsymbol{z},\boldsymbol{z})=\sum_{\alpha\beta}V_{\alpha\beta}z_\alpha z_\beta\geqslant 0\quad (\boldsymbol{z}\in R_0^A).$$

(注意右边实质上也是有限和.) 若令

$$\varphi(\boldsymbol{z})=\exp\left\{i(\boldsymbol{z},\boldsymbol{M})-\frac{1}{2}(\boldsymbol{V}\boldsymbol{z},\boldsymbol{z})\right\},$$

由于 $\varphi(\boldsymbol{z})$ 的任意截口是正态分布 (包含退化的情形) 的特征函数, 因而由

上面改述的 Kolmogoroff 定理, $\varphi(\boldsymbol{z})$ 可以定出 $\mathbb{R}^A(\boldsymbol{B}^A)$ 上的分布. 此分布叫做 $\mathbb{R}^A$ 上的正态分布 $N(\cdot;\boldsymbol{M},\boldsymbol{V})$. 因为此分布在 $(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 上的射影对应于 $\varphi(\boldsymbol{z})$ 在 $(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 上的截口, 故得 $N(\cdot;(M_{\alpha_i}),(V_{\alpha_i\alpha_j}))$. 由此容易知道, $\boldsymbol{M}$ 和 $\boldsymbol{V}$ 分别是此分布的均值和方差.

§5 随机过程

随机过程(stochastic process, random process)是把随时间变动的偶然量抽象起来的概念. 从测度论观点下的概率论来看, 就可这样来定义: 令 $\Omega(\boldsymbol{B},P)$ 为基本的概率空间, T 为实数的集合, 对应于 T 的元素 t 的随机变量 $x_t(\omega)$ 所成的族就叫做随机过程. 从应用上来说, t 表示时间, 而 $x_t(\omega)$ 就表示该偶然量在时间 t 的值. 至于 T 的取法, 既可以是 $\{1,2,3,\cdots\},\{\cdots,-3,-2,-1,0,1,2,3,\cdots\}$ 这样的离散集合, 又可以是区间 $(0,\infty),(-\infty,\infty),(a,b)$ 这样的连续集合. T 是离散集合的情形特别叫做**随机序列**(random sequence).

随机过程 $x_t(\omega),t\in T$, 又可以考虑为无限维的随机向量 $\boldsymbol{x}(\omega)\equiv\Pi_t x_t(\omega)$. $x(\omega)$ 虽是取值于 $\mathbb{R}^T(\boldsymbol{B}^T)$ 内的随机向量, 但因 $\mathbb{R}^T$ 的元素是定义于 T 上的实函数, 所以对于每一 ω 就有一个这样的函数与之对应. 在这种意义下, 随机过程有时也叫做**随机函数**(random function). 与 ω 对应的 t 的函数叫做**样本函数**(sample function) 或者样本过程 (sample process). 因为随机过程是取值于 $\mathbb{R}^T(\boldsymbol{B}^T)$ 内的随机向量, 所以其分布就为 $\mathbb{R}^T(\boldsymbol{B}^T)$ 上的概率分布. 若其为正态分布, 就叫做**正态随机过程**(normal stochastic process).

若 $x_t(\omega)$ 是取复数值的 ω 的可测函数, 则 $x_t(\omega),t\in T$, 叫做复随机过程. 至于把正态随机过程推广到复随机过程的场合, 将在 §27 中讨论.

第2章　可加过程

§6　可加过程的定义

可加过程(additive process, differential process) 是这样的随机过程, 当时间变动时, 它将加上独立的增量. 详细地来说, 对于 $-\infty < a < b \leqslant \infty$, 随机过程 $x_t(\omega)(a \leqslant t < b)$ 叫做可加过程, 如果满足:

(i) $x_a(\omega) \equiv 0$,

(ii) 对 $a \leqslant t_0 < t_1 \cdots < t_n < b$, $x_{t_i} - x_{t_{i-1}}(i = 1, 2, \cdots, n)$是独立的.

同样又可以定义**可加序列**(additive sequence) $x_n(\omega), n = 0, 1, 2, \cdots$, 这时代替 (ii) 的是 $y_n = x_n - x_{n-1}(n = 1, 2, \cdots)$ 为独立的.

下述的**构成定理**是讨论可加序列的基础.

定理 1　对任意的一维分布序列 $\varPhi_1, \varPhi_2, \cdots$, 都可以在适当的概率空间上定义可加序列 $\{x_n\}$, 使得 $x_n - x_{n-1}$ 的分布为 $\varPhi_n$.

证明　假使已经得到了所求的 x_n, 那么, $y_n = x_n - x_{n-1}, n = 1, 2, \cdots$, 就是独立的, 且以 $\varPhi_n$ 为其分布. 故无限维随机向量 $\boldsymbol{y}(\omega) = \prod_i y_i(\omega)$ 的分布就是 $\varPhi_n, n = 1, 2, \cdots$ 的直积分布. 由此可见, 定理可以这样来证明: 设 $T = \{1, 2, \cdots\}, \varOmega = \mathbb{R}^T, \boldsymbol{B} = \boldsymbol{B}^T, P = \prod_n \varPhi_n$, 又对 $\omega \in \varOmega$ 设

$$y_n(\omega) = p_n(\omega), \quad x_n(\omega) = \sum_1^n y_\nu(\omega), \tag{6.1}$$

则 $\{x_n\}$ 就是所求的序列.

下面是关于可加过程的一个相应的定理. 令 $x_t(a \leqslant t < b)$ 为可加过程, $\varPhi_{st}$ 为 $x_t - x_s(s < t)$ 的分布. 因为当 $s < t < u$ 时, $x_u - x_s$ 是独立随机变量 $x_t - x_s$ 和 $x_u - x_t$ 的和, 所以

$$\varPhi_{su} = \varPhi_{st} * \varPhi_{tu} \quad (s < t < u). \tag{6.2}$$

定理 2　若 $\varPhi_{st}(a \leqslant t < b)$ 为满足 (6.2) 的分布系列, 则可以在适当的概率空间上确定可加过程 $x_t(a \leqslant t < b)$, 使得 $x_t - x_s$ 的分布为 $\varPhi_{st}$.

证明 为了找到构造出 x_t 的线索, 假设已经得到了 x_t, 而去求随机向量 $\boldsymbol{x}(\omega)=\prod_t x_t(\omega)$ 的分布的特征函数 φ, 设 $T=[a,b)$. 对 $z\in\mathbb{R}_0^T$,

$$\varphi(z)=E(\mathrm{e}^{i(\boldsymbol{z},\boldsymbol{x})}). \tag{6.3}$$

特别地, 当 z 的坐标除 $z_{t_1},z_{t_2},\cdots,z_{t_n}(t_1<t_2<\cdots<t_n)$ 以外均为 0 时,

$$\begin{aligned}\varphi(z)&=E\left(\mathrm{e}^{i\sum\limits_\nu z_{t_\nu}x_{t_\nu}}\right)\\&=E\left(\mathrm{e}^{i\sum\limits_\nu z'_\nu(x_{t_\nu}-x_{t_{\nu-1}})}\right)\quad\left(z'_\nu=\sum_{i=\nu}^n z_{t_i},t_0=0\right),\end{aligned}$$

由 $\{x_{t_\nu}-x_{t_{\nu-1}}\}$ 的独立性可知上式等于 $\prod\limits_\nu\varphi_{t_{\nu-1}t_\nu}(z'_\nu)$ $(\varphi_{st}=\varPhi_{st}$ 的特征函数), 因而

$$\varphi(\boldsymbol{z})=\prod_\nu\varphi_{t_{\nu-1}}t_\nu(z_{t_\nu}+z_{t_{\nu+1}}+\cdots+z_{t_n}). \tag{6.4}$$

这样一来, 就知道了定义 $\boldsymbol{x}(\omega)$ 的分布的方法, 由此可将定理如下证明. 对 $\boldsymbol{z}\in\mathbb{R}_0^T$, 按照 (6.4) 来定义 $\varphi(\boldsymbol{z})$. 这里必需注意:"比如 $z_{t_1},z_{t_2},\cdots,z_{t_n}$ 中的 z_{t_ν} 等于 0 时, $\varphi(z)$ 虽然也可以由$\varphi_{t_0t_1},\varphi_{t_1t_2},\cdots,\varphi_{t_{\nu-2}t_{\nu-1}}\varphi_{t_{\nu-1}t_{\nu+1}}\varphi_{t_{\nu+1}t_{\nu+2}}\cdots\varphi_{t_{n-1}t_n}$ 来定义, 但要证明它与 (6.4) 的右边一致, $\varphi(z)$ 的定义才能肯定." 为此, 只要证明 $\varphi_{t_{\nu-1}t_{\nu+1}}(\boldsymbol{z})=\varphi_{t_{\nu-1}t_\nu}(\boldsymbol{z})\varphi_{t_\nu t_{\nu+1}}(\boldsymbol{z})$ 就够了, 而由 (6.2) 即知此事必然成立. 下面将证明, 把 (6.4) 的右边看做 $z_{t_1},\cdots,z_{t_n}$ 的函数时, 那就是某个 n 维分布的特征函数. 为此, 考虑使得直积分布 $\prod\limits_\nu\varPhi_{t_{\nu-1}t_\nu}$ 分布于 $R^n(\boldsymbol{B}^n)$ 上的概率空间 $\varOmega'(\boldsymbol{B}',P')$, 且在其上设 $y'_\nu(\omega')=p_\nu(\omega')$ 时, 则其分布为 $\varPhi_{t_{\nu-1}t_\nu}$, 并且这些随机变量都是独立的. 因此, 若设 $x'_\nu(\omega')=\sum\limits_1^\nu y'_i(\omega')$, 且令随机向量 $\boldsymbol{x}'(\omega')=\Pi_\nu x'_\nu(\omega')$ 的分布的特征函数为 $\varphi'(z_1,z_2,\cdots,z_n)$, 则知道 (6.4) 的右边等于 $\varphi'(z_{t_1},z_{t_2},\cdots,z_{t_n})$. 故由 Kolmogoroff 定理 (变形), 得 $\varphi(\boldsymbol{z})$ 是 $R^T(\boldsymbol{B}^T)$ 上的某个分布 (P) 的特征函数. 若设 $\varOmega=\mathbb{R}^T,\boldsymbol{B}=\boldsymbol{B}^T$, 则 $\varOmega(\boldsymbol{B},P)$ 是基本的概率空间, 又若对 $\omega\in\varOmega$, 定义 $x_t(\omega)=p_t(\omega)$, 则 $\{x_t\}$ 就是所求的可加过程.

§7 可加过程的例子

设反复地掷一枚硬币, 令 $x_n(x_0=0)$ 为到第 n 次为止掷出正面的次

数, 若设 $y_1 = x_1, y_2 = x_2 - x_1, y_3 = x_3 - x_2, \cdots$, 则 y_n 随着第 n 次掷出正面或者反面而取 1 或 0. 因此, y_n 的分布 $\varPhi_n$ 都是在 0,1 之上各分布为 1/2 的纯粹不连续分布. 并且 $\{y_n\}$ 是独立的随机变量序列. 故 $\{x_n\}$ 是可加序列, 而其数学模型就可由前节的定理 1 来构成.

下面来考虑随机游动 (random walk) 的问题. 假设一酒醉者从贯通着东西的马路上的一点出发, 一步一步地向东或者向西跑, 这完全是偶然选择的, 设他在走了 n 步后所处的位置为 x_n(在出发点的东边为正, 西边为负). 若设 $y_1 = x_1, y_2 = x_2 - x_1, \cdots$, 则 y_n 随着第 n 步向东跑或者向西跑而取 ± 1, 并且它们的概率分别为 1/2. 也就是说, y_n 的分布 $\varPhi_n$ 就是 $\varPhi_n(\pm 1) = 1/2$ 的纯粹不连续分布. 又因 $\{y_n\}$ 可以考虑为独立随机变量序列, 所以 $\{x_n\}$ 是可加序列, 而这也可由前节的定理 1 来构成.

其次, 以 Poisson 过程 (Poisson process) 作为可加过程的例子. Poisson 过程 $x_t, 0 \leqslant t < \infty$, 是一可加过程, 而且 $x_t - x_s (t < s)$ 的分布 $\varPhi_{st}$ 就等于 Poisson 分布 $P(\cdot; \lambda(t-s))$. 此处 λ 是正的常数. 因为 $\varPhi_{st}$ 满足前节的定理 2 的条件 (6.3) (见 §6), 所以可以利用这个定理. 当 $h \to 0$ 时, 由 Poisson 分布的定义就立即知道

$$P\{\omega / x_{t+h} - x_t \geqslant 1\} = P\{\omega / x_{t+h} - x_t = 1\} \sim \lambda \cdot h.$$

假设某些现象在各个瞬间独立地发生或不发生, 而且令其在 $(t, t + \mathrm{d}t)$ 之间发生的概率为 $\lambda \cdot \mathrm{d}t$. 这时在 $[0, t]$ 之间发生的次数 x_t 可以看做 Poisson 过程 ①.

作为可加过程的另外的一个例子, 可以举出 **Wiener 过程**(Wiener Process). 这其实是随机游动问题的一种连续情形. 它可用一个可加过程 $x_t (0 \leqslant t < \infty)$ 来定义, 使得 $x_t - x_s$ 的分布 $\varPhi_{st}$ 是正态分布 $N(\cdot; 0, t-s)$. 因为 $\varPhi_{st}$ 也满足前节定理 2 的条件 (6.2)(见 §6), 所以可以利用这个定理.

若 x_t 是可加过程, 则 $\alpha x_t + \beta t + \gamma$ 也是可加过程. 又若 $x_t (a \leqslant t < b)$ 和 $y_t (a \leqslant t < b)$ 都是可加过程, 并且随机向量 $\boldsymbol{x} = \prod_t x_t$ 和 $\boldsymbol{y} = \prod_t y_t$ 是独立的 (也可以这样说,“ 若对任意的 $t_1, t_2, \cdots, t_n, n$ 维随机向量 $\prod_i x_{t_i}$ 和

① 这一结论的一个简单证明, 可参见复旦大学数学系郑绍濂等编的《概率论与数理统计》, 上海科学技术出版社, 1961, §6 之例 7. —— 校者注

$\prod_i y_{t_i}$ 是独立的"), 则 $\alpha x_t+\beta y_t(a\leqslant t<b)$ 也是可加过程. 对于两个以上的可加过程来说这一结论也成立. 利用这个事实, 由 Poisson 过程和 Wiener 过程出发, 就可以构成更一般的可加过程 (见 §17).

§8 关于独立随机变量之和的不等式

定理 1(Kolmogoroff 不等式) 设 $x_1,x_2,\cdots,x_n$ 为独立, 且

$$E(x_i)=0,\quad V(x_i)<\infty,\quad i=1,2,\cdots,n, \tag{8.1}$$

则

$$P\left\{\omega/\max_{k=1}^{n}|x_1+\cdots+x_k|\geqslant c\right\}\leqslant\frac{1}{c^2}\sum_{k=1}^{n}V(x_k). \tag{8.2}$$

证明 设 $A_k=\{\omega/|x_1|,|x_1+x_2|,\cdots,|x_1+\cdots+x_{k-1}|<c,|x_1+\cdots+x_k|\geqslant c\},1\leqslant k\leqslant n$, 并令 A_k 的示性函数为 $a_k(\omega).a_k(\omega)$ 关于 $x_1,x_2,\cdots,x_k$ 是可测的. 由 A_k 的定义, 可知 $A_1,A_2,\cdots,A_n$ 互不相交, 于是 (8.2) 的左边等于

$$P(A_1+A_2+\cdots+A_n)=\sum p(A_k).$$

由 $E(x_i)=0$, 得

$$V(x_1+\cdots+x_n)=E((x_1+\cdots+x_n)^2),$$

又因 $\sum a_k\leqslant 1$, 所以上式 $\geqslant\sum_k E(a_k(x_1+\cdots+x_n)^2)$.

但因

$$\begin{aligned}E(a_k(x_1+\cdots+x_n)^2)=&E(a_k(x_1+\cdots+x_k)^2)\\&+2E(a_k(x_1+\cdots+x_k)(x_{k+1}+\cdots+x_n))\\&+E(a_k(x_{k+1}+\cdots+x_n)^2).\end{aligned}$$

由 a_k 的定义可知第 1 项 $\geqslant c^2P(A_k)$. 因 $a_k(x_1+\cdots+x_k)$ 和 $(x_{k+1}+\cdots+x_n)$ 分别关于 $(x_1,\cdots,x_k)$ 和 $(x_{k+1},\cdots,x_n)$ 是可测的, 所以这些随机变量是独立的, 并且 $E(x_{k+1}+\cdots+x_n)=0$, 故第 2 项为 0, 第 3 项 $\geqslant 0$. 因此, 上式 $\geqslant c^2P(A_k)$, 于是

$$V(x_1+\cdots+x_n)\geqslant c^2\sum P(A_k).$$

又由 $\{x_k\}$ 的独立性可知左边等于 $\sum V(x_k)$. 综合以上所述就得 (8.2).

定理 2(Ottaviani 不等式) 若 $x_1,x_2,\cdots,x_n$ 为独立, 并且

$$P\{\omega/|x_{k+1}+\cdots+x_n|\leqslant c\}\geqslant 1/2,\quad k=0,1,\cdots,n-1, \tag{8.3}$$

则

$$P\{\omega/\max_{k=1}^{n}|x_1+\cdots+x_k|>2c\}\leqslant 2P\{\omega/|x_1+\cdots+x_n|>c\}. \tag{8.4}$$

证明 若令

$$\begin{aligned}A_k=&\{\omega/|x_1|,|x_1+x_2|,\cdots,|x_1+\cdots+x_{k-1}|\leqslant 2c,\\&|x_1+\cdots+x_k|>2c\},\quad 1\leqslant k\leqslant n,\\B_k=&\{\omega/|x_{k+1}+\cdots+x_n|\leqslant c\},\quad 1\leqslant k\leqslant n-1,\end{aligned}$$

则 $A_1,A_2,\cdots,A_n$ 互不相交, 于是 (8.4) 左边的 ω 集合等于 $A_1+A_2+\cdots+A_n$. 因由 $|x_1+\cdots+x_k|>2c$, $|x_{k+1}+\cdots+x_n|\leqslant c$ 可以得出 $|x_1+\cdots+x_n|>c$, 所以若令 (8.4) 右边的 ω 集合为 C, 则得

$$A_1\cdot B_1+A_2\cdot B_2+\cdots+A_n\cdot B_n\subseteq C,\quad 此处B_n=\Omega.$$

因为 A_k 由形状为 $(x_1,\cdots,x_k)\in E$ 的条件而定, B_k 由形状为 $(x_{k+1},\cdots,x_n)\in E'$ 的条件而定, 又因 $(x_1,\cdots,x_k)$ 和 $(x_{k+1},\cdots,x_n)$ 是独立的, 于是 $P(A_kB_k)=P(A_k)P(B_k)$. 由 (8.3) 得 $P(B_k)\geqslant 1/2$. 故 $P(A_kB_k)\geqslant P(A_k)/2$. 因此

$$\begin{aligned}P(C)&\geqslant P(A_1B_1)+P(A_2B_2)+\cdots+P(A_nB_n)\\&\geqslant\frac{1}{2}(P(A_1)+P(A_2)+\cdots+P(A_n))\end{aligned}$$

故

$$P(C)\geqslant\frac{1}{2}P(A_1+A_2+\cdots+A_n).$$

§9 0-1 律

当 ω 集合 A 的示性函数 $a(\omega)$ 关于随机向量 $\boldsymbol{x}(\omega)=\prod\limits_{\lambda\in\Lambda}x_\lambda(\omega)$ 为可测时, 就称 A 关于 $\boldsymbol{x}(\omega)$(或关于 $x_\lambda,\lambda\in\Lambda$) 是可测的. 令 λ 为 Λ 的任意

元素, ξ 为任意的实数, 考虑所有形状为 $\{\omega/x_\lambda \leqslant \xi\}$ 的 ω 集合, 而把包含这集合的最小 Borel 集合体记为 $\boldsymbol{B}(\boldsymbol{x})$ 或者 $\boldsymbol{B}(x_\lambda, \lambda \in \Lambda)$. " A 关于 $\boldsymbol{x}(\omega)$ 可测" 就等价于 A 是 $\boldsymbol{B}(\boldsymbol{x})$ 的元素. 换言之, A 是形状为 $\{\omega/\boldsymbol{x}(\omega) \in E\}(E \in \boldsymbol{B}^\Lambda)$ 的集合. 此时, 对于任意的 $\lambda_1, \lambda_2, \cdots, \lambda_n \in \Lambda$ 及 $E \in \boldsymbol{B}^n$ 来说, $\{\omega/(x_{\lambda_1}, x_{\lambda_2}, \cdots, x_{\lambda_n}) \in E\}$ 关于 $\boldsymbol{x}$ 是可测的. 又若 $\lambda_1, \lambda_2, \cdots, \lambda_n \in \Lambda$, 则使得 $x_{\lambda_n}(\omega), n = 1, 2, \cdots$, 收敛的 ω 全体的集合关于 $\boldsymbol{x}$ 也是可测的.

引理 1 令 A 关于 $x_\lambda(\lambda \in \Lambda)$ 为可测. 则对于任意的 $\varepsilon > 0$, 可在 Λ 中找出有限个元素, 例如 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 及存在关于 $x_{\lambda_1}, x_{\lambda_2}, \cdots, x_{\lambda_n}$ 为可测的集合 A_ε, 使得

$$P(A - A_\varepsilon) + P(A_\varepsilon - A) < \varepsilon. \tag{9.1}$$

证明 因为具有上述性质的 A 全体是一 Borel 集合体. 此外, 具有形状为 $\{\omega/x\lambda \leqslant \xi\}$ 的集合具有这种性质, 所以 $\boldsymbol{B}(x)$ 中的元素都具有这种性质.

称 $\boldsymbol{A}_\lambda(\lambda \in \Lambda)$ 是独立的, 意指其示性函数的集合 $\alpha_\lambda(\omega)(\lambda \in \Lambda)$ 构成独立随机变量系.

引理 2 若 $A_1, A_2, \cdots$(有限个或者可数无限个) 是独立的, 则

$$P\Big(\bigcap_n A_n'\Big) = \prod_n P(A_n'), \quad \text{此处} A_n' = A_n \text{或者} A_n^c. \tag{9.2}$$

引理 3 令 $\boldsymbol{x}_\lambda(\omega)(\lambda \in \Lambda)$ 为独立随机向量系, 若对各个 $\lambda \in \Lambda, A_\lambda$ 分别关于 $\boldsymbol{x}_\lambda(\omega)$ 为可测时, 则 $A_\lambda(\lambda \in \Lambda)$ 是独立的.

定理 1(Kolmogoroff 0-1 律) 设 $x_1, x_2, \cdots$ 是独立的, 且若对任意的 n 来说, A 关于 $x_n, x_{n+1}, \cdots$ 为可测, 则

$$P(A) = 0 \text{或者} 1. \tag{9.3}$$

证明 因为 A 关于 $x_1, x_2, \cdots$ 是可测的, 所以对于 ε, 取足够大的 n, 就在 (9.1) 的意义下, 可用关于 $x_1, x_2, \cdots, x_n$ 为可测的 A_ε 来逼近 A. 因为 A 关于 $x_{n+1}, x_{n+2}, \cdots$ 也是可测的, 所以由 $\{x_n\}$ 的独立性, 可知 A 和 A_ε 是相互独立的 (引理 3). 故 $P(A \cdot A_s) = P(A)P(A_s)$. 由 (9.1) 得

$$|P(A) - P(A_\varepsilon)| < \varepsilon, \quad |P(A) - P(A \cdot A_\varepsilon)| < \varepsilon.$$

因此, 令 $\varepsilon \to 0$ 便得 $P(A)=P(A)^2$, 即 $P(A)=0$ 或者 1.

应用定理 1, 就可以证明: 若 $x_1, x_2, \cdots$ 为独立, 则

$$P\Big\{\omega/\sum x_n\text{收敛}\Big\}=0\text{或者}1,$$

$$P\Big\{\omega\Big/\frac{1}{n}\sum_{\nu=1}^{n}x_\nu \to 0\Big\}=0\text{或者}1.$$

定理 2(Borel-Cantelli 定理) 若事件序列 $\{A_n\}$ 满足

$$\sum P(A_n)<\infty, \tag{9.4}$$

则

$$P(\overline{\lim}A_n)=0, \quad P(\underline{\lim}A_n^c)=1. \tag{9.5}$$

若 $\{A_n\}$ 是独立的, 而且

$$\sum P(A_n)=\infty, \tag{9.6}$$

则

$$P(\overline{\lim}A_n)=1, \quad P(\underline{\lim}A_n^c)=0. \tag{9.7}$$

证明 若假定 (9.4) 成立, 则

$$P(\overline{\lim}A_n)=P\Big(\bigcap_k\bigcup_{n\geqslant k}A_n\Big)\leqslant P\Big(\bigcup_{n\geqslant k}A_n\Big)\leqslant\sum_{n\geqslant k}P(A_n)\to 0.$$

故 (9.5) 的第一式即被证明. 利用 $\underline{\lim}A_n^c=(\overline{\lim}A_n)^c$, 第二式就可由第 1 式导出.

其次令 $\{A_n\}$ 是独立的, 且令 (9.6) 成立.

$$P(\underline{\lim}A_n^c)=P\Big(\bigcup_k\bigcap_{n\geqslant k}A_n^c\Big)\leqslant\sum_k P\Big(\bigcap_{n\geqslant k}A_n^c\Big).$$

由引理 2,

$$P\Big(\bigcap_{n\geqslant k}A_n^c\Big)=\prod_{n\geqslant k}P(A_n^c)=\prod_{n\geqslant k}(1-P(A_n)).$$

因由 (9.6) 得 $\sum\limits_{n\geqslant k}P(A_n)=\infty$, 所以上面的无穷乘积为 0, 于是 $P(\underline{\lim}A_n^c)$ 也为 0. 因而 $P(\overline{\lim}A_n)=1$.

§10 可加序列的收敛

令 $x_n(n=0,1,2,\cdots)$ 为可加序列. 若设 $y_n = x_n - x_{n-1}$, 则 $\{y_n\}$ 是独立的随机变量序列. 本节的目的是来考察当 $n\to\infty$ 时 x_n 的收敛条件. 若以 C 表示使 $x_n(\omega)(n=0,1,2,\cdots)$ 收敛的 ω 集合, 则得

$$C=\bigcap_p\bigcup_k\bigcap_{m,n>k}\{\omega/|x_m(\omega)-x_n(\omega)|<1/p\}$$
$$=\bigcap_p\bigcup_k\bigcap_{m,n>k}\{\omega/|y_{n+1}(\omega)+\cdots+y_m(\omega)|<1/p\}.$$

故 C 关于 $y_1,y_2,\cdots$ 是可测的. 因为 $x_n(\omega)(n\geqslant 0)$ 的收敛等价于 $x_n(\omega)-x_N(\omega)(n\geqslant N)$ 的收敛, 所以也可以写成

$$C=\bigcap_p\bigcup_{k\geqslant N}\bigcap_{m,n>k}\{\omega/|y_{n+1}(\omega)+\cdots+y_m(\omega)|<1/p\},$$

而 C 又可以说是关于 $y_{N+1},y_{N+2},\cdots$ 是可测的. 故由 Kolmogoroff 0-1 律, 就得

$$P(C)=0\text{或者}1. \tag{10.1}$$

那么在什么样的条件下, $P(C)$ 才为 1 呢? 首先叙述充分条件.

定理 1 若 $\{E(x_n)\}$ 与 $\{V(x_n)\}$ 同时收敛, 则 $P(C)=1$, 也就是说, x_n 几乎处处收敛.

证明 由于 $x_n=(x_n-E(x_n))+E(x_n), V(x_n)=V(x_n-E(x_n))$, 因而设 $E(x_n)=0$ 也不会丧失普遍性. 因为

$$\sum_1^\infty V(y_n)=\lim_{n\to\infty}V(x_n)<\infty,$$

所以对任意的 $\varepsilon>0$, 存在 $n(\varepsilon)$, 而且当 $n\geqslant n(\varepsilon)$ 时, 总有

$$\sum_n^\infty V(y_n)<\varepsilon.$$

根据 Kolmogoroff 不等式, 得

$$P\left\{\omega/\max_{1\leqslant k\leqslant m}|y_{n+1}+\cdots+y_{n+k}|>a\right\}\leqslant\varepsilon/a^2.$$

令 $m \to \infty$ 便得

$$P\Big\{\omega/\sup_k |y_{n+1}+\cdots+y_{n+k}| > a\Big\} \leqslant \varepsilon/a^2.$$

因为 $|y_{n+k}+\cdots+y_{n+l}| \leqslant |y_{n+1}+\cdots+y_{n+k-1}|+|y_{n+1}+\cdots+y_{n+l}|$ $(k<l)$, 所以由上式得出

$$P\Big\{\omega/\sup_{l>k} |y_{n+k}+\cdots+y_{n+l}| > 2a\Big\} \leqslant \varepsilon/a^2 \quad (n \geqslant n(\varepsilon)),$$

令 $n \to \infty$ 就得

$$P\Big\{\omega/\lim_{n\to\infty}\sup_{l>k} |y_{n+k}+\cdots+y_{n+l}| > 2a\Big\} \leqslant \varepsilon/a^2,$$

先令 $\varepsilon \to 0$, 而后再令 $a \to 0$ 就得

$$P\Big\{\omega/\lim_{n\to\infty}\sup_{l>k} |y_{n+k}+\cdots+y_{n+l}| > 0\Big\} = 0.$$

这就说明了 C 的余集的 P 测度为 0.

下述定理说明了几乎处处收敛的必要而且充分条件.

定理 2(三级数定理) 若设当 $|y_n(\omega)| \leqslant 1$ 时 $y'_n(\omega) = y_n(\omega)$, 当 $|y_n(\omega)| > 1$ 时 $y'_n(\omega) = 0$, 则 x_n 几乎处处收敛的充分必要条件是下面的三个级数都收敛:

$$\sum E(y'_n), \quad \sum V(y'_n), \quad \sum P\{\omega/y_n \neq y'_n\}. \tag{10.2}$$

证明 充分性. 由于前两个级数的收敛性, 故根据定理 1 可知 $\sum y'_n$ 是几乎处处收敛的. 由第 3 个级数的收敛性, 而且根据 Borel-Cantelli 定理, 可知在某项之后 $y_n(\omega) = y'_n(\omega)$ 的概率为 1. 故由 y'_n 的几乎处处收敛性, 而得出 y_n 是几乎处处收敛的.

必要性. 若设 $\sum P\{\omega/y_n \neq y'_n\} = \infty$, 则根据 Borel-Cantelli 定理, 可知对无限多个 $n, y_n(\omega) \neq y'_n(\omega)$ 的概率就是 1. 因为 $y_n \neq y'_n$ 等价于 $|y_n| > 1$, 所以无限多个 y_n 的绝对值大于 1 的概率也是 1. 故 $\sum y_n$ 发散的概率就为 1. 所以为了使 $\sum y_n$ 即 x_n 几乎处处收敛, 就需要 $\sum P\{\omega/y_n \neq y'_n\} < \infty$. 若此式成立, 则再次根据 Borel-Cantelli 定理, 可知由 $\sum y_n$ 的几乎处处收敛性就会得出 $\sum y'_n$ 的几乎处处收敛性. 现在令 y'_n 的分布为

Φ_n, 而令 Φ_n 的负反分布为 $\check{\Phi}_n$, 又令 $\Phi_n * \check{\Phi}_n$ 为 $\widetilde{\Phi}_n$. 因为 Φ_n 的支集为 $[-1,1]$, 所以 $\widetilde{\Phi}_n$ 的支集就为 $[-2,2]$. 显然有

$$M(\widetilde{\Phi}_n)=0, \quad V(\widetilde{\Phi}_n)=2V(\Phi_n)=2V(y_n').$$

若令 Φ_n 的特征函数为 $\varphi_n(z)$, 则 $\widetilde{\Phi}_n$ 的特征函数就为 $|\varphi_n(z)|^2$. 如上面所证, 因为 $\sum y_n'$ 是几乎处处收敛的, 所以

$$\lim_{n\to\infty} E(\mathrm{e}^{iz(y_1'+\cdots+y_n')})=E(\mathrm{e}^{izx_\infty}), \quad x_\infty=\sum y_n=\lim x_n,$$

因此, $\prod\limits_n \varphi_n(z)$ 在 $z=0$ 的某些邻域 $(|z|<a)$ 上不为 0, $\prod\limits_n |\varphi_n(z)|^2 > 0(|z|<a)$, 由此得出

$$\sum_n (1-|\varphi_n(z)|^2)<\infty,$$

也就是

$$\sum_n \int (1-\cos z\xi)\widetilde{\Phi}_n(\mathrm{d}\xi)<\infty.$$

因为对足够小的 ξ 来说, 有 $1-\cos\xi > \xi^2/3$, 所以若令 z 为足够小的正数 (注意 Φ_n 的支集是 $[-2,2]$), 则得

$$\sum_n \int \frac{z^2\xi^2}{3}\widetilde{\Phi}_n(\mathrm{d}\xi)<\infty.$$

由此可知 $\sum V(\widetilde{\Phi}_n)$ 收敛, 从而 $\sum V(y_n')$ 是收敛的. 故由定理 1 可知 $\sum_n (y_n' - E(y_n'))$ 是几乎处处收敛的. 但因 $\sum\limits_n y_n'$ 由假设是几乎处处收敛的, 所以知 $\sum\limits_n E(y_n')$ 也是收敛的.

定理 3　下述的 3 个条件是等价的:

(i) x_n 的分布收敛,

(ii) x_n 依概率收敛,

(iii) x_n 几乎处处收敛.

证明　(iii) ⇒ (ii)⇒ (i) 是显然的. 先证 (i)⇒(ii). 若令 y_n 的分布为 Φ_n, 其特征函数为 φ_n, 则由 (i) 知 $\Phi_1 * \Phi_2 * \cdots * \Phi_n$ 当 $n\to\infty$ 时收敛于某一分布 Φ. 若令 Φ 的特征函数为 φ 时, 则 $|\varphi(z)|$ 在 $z=0$ 的某些邻域

$|z| \leqslant a$ 上大于某一正数 b. 又因 $\varphi_1(z)\cdot\varphi_2(z)\cdot\cdots$ 广义地一致收敛于 $\varphi(z)$, 所以对任意的 $\varepsilon>0$, 可以定出 $N(\varepsilon)$, 使得

$$m>n>N(\varepsilon) \Rightarrow |\varphi_{n+1}(z)\cdot\varphi_{n+2}(z)\cdots\varphi_m(z)-1|<s \quad (|z|\leqslant a).$$

现在若置 $\theta(z)=\varphi_{n+1}(z)\cdots\varphi_m(z)$, 且令对应于 $\theta(z)$ 的分布即 x_m-x_n 的分布为 Θ, 则当 $m>n>N(\varepsilon)$ 时就有

$$|\Theta(z)-1|<\varepsilon \quad (|z|\leqslant a),$$

故

$$\left|\int(1-\mathrm{e}^{izx})\Theta(\mathrm{d}x)\right|<\varepsilon.$$

若从 $-a$ 到 a 对 z 积分, 然后再除以 $2a$, 则得

$$\int\left(1-\frac{\sin xa}{xa}\right)\Theta(\mathrm{d}x)<\varepsilon.$$

因存在常数 $C>0$, 使得

$$1-\frac{\sin x}{x}\geqslant C\frac{x^2}{1+x^2},$$

所以

$$\int\frac{x^2a^2}{1+x^2a^2}\Theta(\mathrm{d}x)<\frac{\varepsilon}{C},$$

因而

$$\int_{|x|\geqslant\eta}\Theta(\mathrm{d}x)<\frac{(1+\eta^2a^2)\varepsilon}{C\eta^2a^2},$$

综上所述, 当 $m>n>N(\varepsilon)$ 时, 有

$$P\{\omega/|x_m-x_n|\geqslant\eta\}<(1+\eta^2a^2)\varepsilon/C\eta^2a^2.$$

这就说明了 x_n 是依概率收敛的. 剩下的就是证明 (ii)⇒(iii), 但这与定理 1 的证明方法相似. 只要以 Ottaviani 不等式来代替 Kolmogoroff 不等式就行了.

§11 散 布 度

如前面所述, 方差是表示分布散布程度的标志, 但方差不一定存在, 它的运用范围就有所限制. 因此引进散布度这样一个概念, 散布度对所有

的分布都存在而且起着与方差同样的作用. 对一维分布 Φ, 就称下式中的 $\delta(\Phi)$ 为 Φ 的**散布度**:

$$\delta(\Phi) = -\ln\left[\iint \mathrm{e}^{-|x-y|}\Phi(\mathrm{d}x)\Phi(\mathrm{d}y)\right]. \tag{11.1}$$

再把上式括号里的二重积分的值表示为 $q(\Phi)$, 并称它为 Φ 的**集中度**. 对随机变量 x, 也可将 $\delta(x)$ 和 $q(x)$ 分别定义为由 x 的分布 Φ 所决定的 $\delta(\Phi)$ 和 $q(\Phi)$. 由定义知道下面结论成立.

(i) $0 < q(\Phi) \leqslant 1, 0 \leqslant \delta(\Phi) < \infty$.

(ii) $q(x) = q(x+a) = q(-x), \delta(x) = \delta(x+a) = \delta(-x)$.

(iii) $q(\Phi) = 1 \Leftrightarrow \delta(\Phi) = 0 \Leftrightarrow \Phi$ 是 δ 分布.

(iv) $\delta(x_n) \to 0 \Leftrightarrow$ 存在 $\{a_n\}$, 使得 $x_n - a_n$ 依概率收敛于 0.

证明 $\Leftarrow$ 是显然的. 再来证明 $\Rightarrow$. 由 $\delta(x_n) \to 0$ 得出

$$\iint \mathrm{e}^{-|x-y|}\Phi_n(\mathrm{d}x)\Phi_n(\mathrm{d}y) \to 1 \quad (\Phi_n\text{是}x_n\text{的分布}).$$

故存在 $\{a_n\}$ 使得

$$\int \mathrm{e}^{-|x-a_n|}\Phi_n(\mathrm{d}x) \to 1,$$

也就是

$$\int (1-\mathrm{e}^{-|x-a_n|})\Phi_n(\mathrm{d}x) \to 0,$$

故得

$$\int_{|x-a_n|>\varepsilon} \Phi_n(\mathrm{d}x) \leqslant \frac{1}{1-\mathrm{e}^{-\varepsilon}} \int (1-\mathrm{e}^{-|x-a_n|})\Phi_n(\mathrm{d}x) \to 0.$$

(v) $\delta(x_n) \to \infty \Leftrightarrow$ 对所有的 $l, Q_n(l) \equiv \sup\limits_a \Phi_n[a-l, a+l] \to 0$.

证明 若令 $\delta(x_n) \to \infty$, 则 $q(x_n) \to 0$.

$$\begin{aligned}\Phi_n[a-l,a+l]^2 &= \iint\limits_{|x-a|,|y-a|\leqslant l} \Phi_n(\mathrm{d}x)\Phi_n(\mathrm{d}y)\\ &\leqslant \iint\limits_{|x-y|\leqslant 2l} \Phi_n(\mathrm{d}x)\Phi_n(\mathrm{d}y) \leqslant \mathrm{e}^l q(x_n),\end{aligned}$$

故得

$$Q_n(l) \leqslant \mathrm{e}^{l/2} q(x_n)^{1/2} \to 0.$$

反之, 对所有的 l, 令 $Q_n(l) \to 0$.

$$\int \mathrm{e}^{-|x-y|} \Phi_n(\mathrm{d}x) \leqslant \int_{|x-y| \geqslant l} '' + \int_{|x-y|<l} '' \leqslant \mathrm{e}^{-l} + Q_n(l),$$

$$q(x_n) = \int\int \mathrm{e}^{-|x-y|} \Phi_n(\mathrm{d}x) \Phi_n(\mathrm{d}y) \leqslant \mathrm{e}^{-l} + Q_n(l).$$

若取 l 足够大, 使 $\mathrm{e}^{-l} < \varepsilon/2$, 再取 n 足够大使 $Q_n(l) < \varepsilon/2$, 则得 $q(x_n) < \varepsilon$, 因而 $q(x_n) \to 0$. 故 $\delta(x_n) \to \infty$.

(vi) 若令 φ 为 Φ 的特征函数, 则

$$q(\Phi) = \frac{1}{\pi} \int \frac{|\varphi(z)|^2}{1+z^2} \mathrm{d}z.$$

因此, 若 $\Phi_n \to \Phi$, 则 $q(\Phi_n) \to q(\Phi)$, 也就是 $\delta(\Phi_n) \to \delta(\Phi)$. 同理, 若 $x_n \to x$(依概率收敛), 则 $\delta(x_n) \to \delta(x)$.

证明 若令 $\Psi = \check{\Phi}$(Φ 的负反), 则 $\Phi * \Psi$ 的特征函数为 $|\varphi(z)|^2$.

$$\begin{aligned} q(\Phi) &= \iint \mathrm{e}^{-|x-y|} \Phi(\mathrm{d}x) \Phi(\mathrm{d}y) = \iint \mathrm{e}^{-|x+y|} \Phi(\mathrm{d}x) \Psi(\mathrm{d}y) \\ &= \int \mathrm{e}^{-|x|} (\Phi * \Psi)(\mathrm{d}x) = \int \frac{1}{\pi} \int \frac{\mathrm{e}^{izx}}{1+z^2} \mathrm{d}z (\Phi * \Psi)(\mathrm{d}x) \\ &= \frac{1}{\pi} \iint \mathrm{e}^{izx} (\Phi * \Psi)(\mathrm{d}x) \frac{\mathrm{d}z}{1+z^2} = \frac{1}{\pi} \int \frac{|\varphi(z)|^2}{1+z^2} \mathrm{d}z. \end{aligned}$$

(vii) **散布度增加原理** 若 x, y 是独立的, 则

$$q(x+y) \leqslant q(x) \quad 即 \quad \delta(x+y) \geqslant \delta(x),$$

而等号只能在 y 的分布是 δ 分布时才成立.

证明 令 x 和 y 的分布的特征函数分别为 φ_1 和 φ_2, 则

$$q(x+y) = \frac{1}{\pi} \int \frac{|\varphi_1(z)\varphi_2(z)|^2}{1+z^2} \mathrm{d}z \leqslant \frac{1}{\pi} \int \frac{|\varphi_1(z)|^2}{1+z^2} \mathrm{d}z = q(x).$$

为使等号成立, 对 $|\varphi_1(z)| > 0$ 的 z 必需有 $|\varphi_2(z)|^2 = 1$. 然而因在 $z = 0$ 的邻域上 $|\varphi_1(z)| > 0$, 所以在同一的邻域上得 $|\varphi_2(z)|^2 = 1$. 令 y 的分布

为 $\varPhi, \varPhi$ 和负反 $\check{\varPhi}$ 的卷积为 $\widetilde{\varPhi}$. 因为 $\widetilde{\varPhi}$ 的特征函数是 $|\varphi_2(z)|^2$, 并因 $\widetilde{\varPhi}$ 是对称分布, 所以当 $|z| \leqslant a$ 时, 有

$$\int \cos(z\xi)\widetilde{\varPhi}(\mathrm{d}\xi) = 1 \quad \text{即} \quad \int (1-\cos(z\xi))\widetilde{\varPhi}(\mathrm{d}\xi) = 0.$$

若从 $-a$ 到 a 对 z 积分, 并以 $2a$ 去除, 则得

$$\int \left(1 - \frac{\sin(a\xi)}{a\xi}\right)\widetilde{\varPhi}(\mathrm{d}\xi) = 0.$$

因为

$$1 - \frac{\sin x}{x} > C\frac{x^2}{1+x^2}, \quad C = \text{正的常数},$$

所以

$$\int \frac{a^2\xi^2}{1+a^2\xi^2}\widetilde{\varPhi}(\mathrm{d}\xi) \leqslant 0,$$

故 $\widetilde{\varPhi}$ 是单位分布 $\delta(\cdot;0)$, 而 $\varPhi$ 就是一般的 δ 分布.

现在令 x_n 为可加序列, 并设 $y_n = x_n - x_{n-1}$. 因为 x_{n+1} 是相互独立的 x_n 和 y_{n+1} 的和, 所以 $q(x_n) \geqslant q(x_{n+1})$. 因而 $\delta(x_n) \leqslant \delta(x_{n+1})$. 若设 $q = \lim q(x_n), \delta = \lim \delta(x_n)$, 则下面的定理成立.

定理 1 $\delta < \infty \Leftrightarrow q > 0 \Leftrightarrow$ 存在 $\{a_n\}$ 使得 $\{x_n - a_n\}$ 几乎处处收敛, 并且 $\delta = \delta(\lim\limits_n(x_n - a_n))$.

定理 2 $\delta = \infty \Leftrightarrow q = 0 \Leftrightarrow Q_n(l) = \sup\limits_a P\{\omega/a \leqslant x_n \leqslant a + l\} \to 0$.

证明 定理 2 早已在 (v) 里叙述过. 再来证明定理 1. 令 $\delta < \infty$, 也就是 $q > 0$.

$$q = \lim_n \frac{1}{\pi}\int \frac{|\varphi_1(z)|^2 \cdots |\varphi_n(z)|^2}{1+z^2}\mathrm{d}z = \frac{1}{\pi}\int \prod_{n=1}^{\infty} |\varphi_n(z)|^2 \frac{\mathrm{d}z}{1+z^2}.$$

由于 $q > 0$, 故可适当地选取一测度为正的集合 A, 使得在其上有 $\prod\limits_n |\varphi_n(z)|^2 > 0$, 因此

$$\sum_n [1 - |\varphi_n(z)|^2] < \infty.$$

故若适当地取一测度为正的集合 A_1, 则能在其上得

$$\sum_n [1 - |\varphi_n(z)|^2] < C < \infty.$$

A_1 的测度 (记为 $|A_1|$) 可以假设为有限的. 若令 y_n 的分布为 Φ_n, Φ_n 和它的负反 $\check{\Phi}_n$ 的卷积为 $\widetilde{\Phi}_n$, 则

$$\sum_n \int (1-\cos(z\xi))\widetilde{\Phi}_n(\mathrm{d}\xi) < C. \tag{11.2}$$

因为 $0 < |A_1| < \infty$, 所以

$$f(\xi) \equiv \frac{1+\xi^2}{\xi^2}\int_{A_1}(1-\cos(z\xi))\mathrm{d}z$$

在 $0 < |\xi| < \infty$ 内为连续, 并且 $f(\xi) > 0$. 当 $|\xi| \to \infty$ 时, 就由 Riemann-Lebesgue 定理 ① 得 $f(\xi) \to |A_1| > 0$. 当 $|\xi| \to 0$ 时, 就得

$$f(\xi) \to \frac{1}{2}\int_{A_1} z^2\mathrm{d}z > 0,$$

故 $f(\xi)$ 的下确界为正, 而且

$$\int_{A_1}(1-\cos(z\xi))\mathrm{d}z > k\frac{\xi^2}{1+\xi^2} \quad (k>0).$$

因此, 将 (11.2) 对 z 在 A_1 上积分, 就得

$$\sum_n \int \frac{\xi^2}{1+\xi^2}\widetilde{\Phi}_n(\mathrm{d}\xi) < C'.$$

因为

$$\int \frac{\xi^2}{1+\xi^2}\widetilde{\Phi}_n(\mathrm{d}\xi) = \int\int \frac{(\xi-\eta)^2}{1+(\xi-\eta)^2}\Phi_n(\mathrm{d}\xi)\Phi_n(\mathrm{d}\eta),$$

所以适当地取 $\{b_n\}$ 就得

$$\sum_n \int \frac{(\xi-b_n)^2}{1+(\xi-b_n)^2}\Phi_n(\mathrm{d}\xi) < \sum_n \int \frac{\xi^2}{1+\xi^2}\widetilde{\Phi}_n(\mathrm{d}\xi) < C'.$$

因而

$$\sum_n \int_{|\xi-b_n|>1}\Phi_n(\mathrm{d}\xi) < \infty, \quad \sum_n \int_{|\xi-b_n|\leqslant 1}\xi^2\Phi_n(\mathrm{d}\xi) < \infty.$$

---

① 参见 S. Bochner 与 K. Chandrasekharan 著的中译本《傅里叶变式》, 何旭初译, 高等教育出版社, 第 1 章 §2. —— 校者注

现在若设 $z_n = y_n - b_n$(当 $|y_n - b_n| \leqslant 1$ 时),=0(当 $|y_n - b_n| > 1$ 时), 则由上面的结果可推得

$$\sum_n P\{\omega/z_n \neq y_n - b_n\} < \infty,$$

$$\sum_n E(z_n^2) < \infty.$$

由第 1 条件, 并且根据 Borel-Cantelli 定理, 就知在某项之后 $z_n = y_n - b_n$ 的概率为 1. 又由第 2 条件而得出

$$\sum_n V(z_n) \leqslant \sum_n E(z_n^2) < \infty,$$

根据 §10 的定理 1, 可知 $\sum\limits_n (z_n - c_n)(c_n = E(z_n))$ 是几乎处处收敛的. 故 $\sum\limits_n (y_n - d_n)(d_n = b_n + c_n)$ 也是几乎处处收敛. 因而 $\{x_n - a_n\}(a_n = d_1 + d_2 + \cdots + d_n)$ 也是几乎处处收敛. 若令 $\lim(x_n - a_n)$ 的特征函数为 $\varphi(z)$, 则因 $|\varphi(z)|^2 = \prod\limits_{n=1}^{\infty} |\varphi_n(z)|^2$, 而得 $q(\lim(x_n - a_n)) = q$, 因而 $\delta(\lim(x_n - a_n)) = \delta$.

上述定理 1 中的数列 $\{a_n\}$ 叫做**收敛化常数列**. 若 $\{a_n\}$ 是收敛化常数列, $\{b_n\}$ 是收敛数列, 则 $\{a_n + b_n\}$ 是收敛化常数列. 作为求收敛化常数列的一个方法, 有下述的定理.

定理 3 若定理 1 的条件得到满足时, 用下列等式

$$E(\arctan(x_n - c_n)) = 0 \qquad (*)$$

决定 c_n, c_n 所成的数列 $\{c_n\}$ 就是收敛化常数列. 这叫做 **Doob 收敛化常数列.**

证明 首先在上述的条件下证明 c_n 是唯一确定的. 若令 c_n 从 $-\infty$ 变到 ∞, 则 $E(\arctan(x_n - c_n))$ 从 π/2(狭义) 单调地连续减少到 −π/2. 故存在一个而且唯一的一个 c_n 使 (*) 式成立. 若定理 1 的条件满足, 则存在收敛化常数列 $\{a_n\}$. 若能证明 $\{c_n - a_n\}$ 是收敛序列, 那么定理 3 就被证明了. 为此, 只要证明 $\{c_n - a_n\}$ 的极限点是唯一的就行了. 现在令

$$c_{p(n)} - a_{p(n)} \to c \quad (p(n)\text{是自然数的子序列}).$$

因为

$$x_{p(n)} - c_{p(n)} = (x_{p(n)} - a_{p(n)}) - (c_{p(n)} - a_{p(n)}) \to x - c,$$

所以

$$E(\arctan(x-c)) = \lim_n E(\arctan(x_{p(n)} - c_{p(n)})) = 0.$$

故 c 是唯一确定的.

注意 在定理 1 中, $\sum y_n$ 叫做**收敛型**, 而在定理 2 中, 就叫做**发散型**.

§12 可加过程的简单性质

令 $x_t, t \in T = [a, b)$ 为可加过程. 由可加过程的定义和散布度增加原理, 可得

$$(s,t) \subseteq (u,v) \Rightarrow \delta(x_v - x_u) \geqslant \delta(x_t - x_s).$$

特别设 $u = s = a$, 就知道 $\delta(t) \equiv \delta(x_t)$ 是 t 的增函数. $\delta(t)$ 的不连续点的集合 D 至多是可数的. D 是下述三集合的和.

$$D^+ = \{t/\delta(t+0) > \delta(t), \delta(t-0) = \delta(t)\},$$
$$D^- = \{t/\delta(t+0) = \delta(t), \delta(t-0) < \delta(t)\},$$
$$D^0 = \{t/\delta(t+0) > \delta(t), \delta(t-0) < \delta(t)\}.$$

如前节所述, 对每一 t, 使得

$$E\{\arctan(x_t - f(t))\} = 0$$

的常数 $f(t)$ 是唯一确定的. 现在设

$$z_t = x_t - f(t), \quad x_t = z_t + f(t).$$

令 $t_1 < t_2 < \cdots \to t$ 时, 因为 $z_{t_n} (n = 1, 2, \cdots)$ 是可加序列并且 $\delta(t_n) \leqslant \delta(t)$, 所以 $\{z_{t_n} - c_n\}$ 几乎处处收敛. 因为 $E(\arctan z_{t_n}) = 0$, 所以可取 $c_n = 0$. 也就是说, $\{z_{t_n}\}$ 是几乎处处收敛的. 其极限记为 z_{t-}. 对于固定的 t, z_{t-} 除了一 P 测度为 0 的集合外可以唯一确定 (例外集合是随 t 而变的). 又若对 $t_1' < t_2' < \cdots$ 来考虑同样的极限, 令它为 z_{t-}', 并把 $\{t_n\}$ 和 $\{t_n'\}$ 凑在一起, 令 $t_1'' < t_2'' < \cdots \to t$, 而再令所对应的极限为 z_{t-}'' 时, 则得

$$z_{t-} = z_{t-}'' = z_{t-}' \quad \text{(a.e.)}.$$

故 z_{t-} 具概率为 1 地与 $\{t_n\}$ 的选择无关. 但例外的 ω 集合是随 t 而变的. 为了定义 z_{t+}, 令 $t_n \downarrow t$, 且来考虑 $\{z_{t_n} - z_{t_1}\}$, 这也是可加序列, 又因 $\delta(z_{t_n} - z_{t_1}) = \delta(z_{t_1} - z_{t_n})\delta(z_{t_1} - z_t)$, 所以 $\{z_{t_n} - z_{t_1} - c_n\}$ 几乎处处收敛. 因此, $\{z_{t_n} - c_n\}$ 是几乎处处收敛的. 因为 $E(\arctan(z_{t_n} - c_n)) = 0$, 所以 $\{z_{t_n}\}$ 几乎处处收敛. 令其极限为 z_{t_+}. 这也与 $\{t_n\}$ 的选择无关.

定理 1

$$t \notin D \Leftrightarrow P\{\omega/z_{t_+} = z_t = z_{t-}\} = 1,$$

$$t \in D^+ \Leftrightarrow P\{\omega/z_{t-} = z_t\} = 1, \quad \delta(z_{t+} - z_t) > 0,$$

$$t \in D^- \Leftrightarrow P\{\omega/z_{t+} = z_t\} = 1, \quad \delta(z_t - z_{t-}) > 0,$$

$$t \in D^0 \Leftrightarrow \delta(z_{t+} - z_t) > 0, \quad \delta(z_t - z_{t-}) > 0.$$

在证明之前, 先证下面的引理.

引理 若 $\{x_n, y_n, \cdots, u_n\}$ 对所有的 $n = 1, 2, \cdots$ 是独立的, 并且 $x_n, y_n, \cdots, u_n$ 当 $n \to \infty$ 时分别几乎处处收敛于 $x, y, \cdots, u$, 则 $x, y, \cdots, u$ 也是独立的.

证明

$$\begin{aligned}
&E\{\exp\{i(\theta x + \varphi y + \cdots + \psi u)\}\} \\
&= \lim_{n\to\infty} E\{\exp\{i(\theta x_n + \varphi y_n + \cdots + \psi u_n)\}\} \\
&= \lim_{n\to\infty} E\{e^{i\theta x_n}\}E\{e^{i\varphi y_n}\}\cdots E\{e^{i\psi u_n}\} \\
&= E\{e^{i\theta x}\}E\{e^{i\varphi y}\}\cdots E\{e^{i\psi u}\}.
\end{aligned}$$

若令 $x, y, \cdots, u$ 的分布为 $\Theta, \Phi, \cdots, \Psi$, 随机向量 $(x, y, \cdots, u)$ 的分布为 F, 则由上式知道, F 的特征函数就是 $\Theta \times \Phi \times \cdots \times \Psi$ 的特征函数, 因而这两个分布相等. 这意味着 $x, y, \cdots, u$ 是独立的.

定理的证明 由上面的引理, 可知 $z_t - z_{t-}$ 和 z_{t-} 是独立的, 并因 $z_t = (z_t - z_{t-}) + z_{t-}$, $\delta(z_t) = \lim\limits_{s\uparrow t} \delta(z_s) = \delta(t-0)$, 所以若 $\delta(t-0) = \delta(t)$, 则 $z_t - z_{t-}$ 为一常数 (具概率为 1). 由于 $E(\arctan z_t) = 0$, 而得 $E(\arctan z_{t-}) = \lim\limits_{s\uparrow t} E(\arctan z_s) = 0$. 所以这常数应该是 0, 也就是说, $P(z_t = z_{t-}) = 1$, 若 $\delta(t-0) < \delta(t)$, 则 $z_t - z_{t-}$ 不为常数, 且得 $\delta(z_t - z_{t-}) > 0$. 对于 z_t 来说也是一样. 综合起来就得上述的定理.

下面将把所有的 z_t 的跃度 $z_{t+}-z_{t-}, t\in D$, 凑起来定义可加过程 u_t. 对这个 u_t, 不允许简单地用公式

$$u_t=\sum_{\substack{s\leqslant t\\ s\in D}}(z_{s+}-z_{s-})$$

作为它的定义, 因为这样发生了无限和的收敛问题. 因为 D 是可数集合, 所以可把其元素排成一列 $s_1, s_2, \cdots$. 设

$$u_t^{(n)}=\sum_{\substack{1\leqslant i\leqslant n\\ s_i\leqslant t}}(z_{s_i+}-z_{s_i-})-c_t^{(n)},$$

但当 $s_i=t$ 时, 就以 z_t 代替 z_{s_i+}. $c_t^{(n)}$ 是使 $E(\arctan u_t^{(n)})=0$ 的常数. 因为 $u_t^{(n)}$ 是独立随机变量的可数和的部分和, 并由散布度增加原理, 得 $\delta(u_t^{(n)})\leqslant\delta(z_t)$, 所以 $u_t^{(n)}$ 是几乎处处收敛的. 此极限可以记为 u_t. 显然 $E(\arctan u_t)=0$. 注意到这点就会知道, 纵使改变 D 的排列也可以得到同样的 u_t. 由定义显然有以下定理.

定理 2 u_t 是可加过程, 而 $\delta(u_t)$ 是仅在 D 中的点上产生跳跃, 且具有正的跃度的纯粹不连续的增函数, 则

$$\text{当 } t\notin D \text{ 时, } P\{\omega/u_{t-}=u_t=u_{t+}\}=1,$$

$$\text{当 } t\in D \text{ 时, } P\{\omega/u_{t+}-u_t=z_{t+}-z_t+\text{ 常数}\}=1,$$

$$P\{\omega/u_t-u_{t-}=z_t-z_{t-}+\text{常数}\}=1.$$

其次设 $v_t=z_t-u_t-c_t$(选 c_t 使得 $E(\arctan v_t)=0$), 而去考察 v_t 的性质.

定理 3 v_t 是可加过程, 而 $\delta(u_t)$ 是 t 的连续函数, 则

$$P\{\omega/v_{t-}=v_t=v_{t+}\}=1. \tag{12.1}$$

证明 若设 $v_t^{(n)}=z_t-u_t^{(n)}-c_t$, 则这显然是可加过程. 因为 $v_t^{(n)}$ 几乎处处收敛于 v_t, 所以利用上面的引理 1, 就知道 v_t 也是可加过程. 由 $E(\arctan v_t)=0$, 就可以确定 v_{t-} 和 v_{t+}. 因此, c_{t-0} 和 c_{t+0} 也可以确定. 由定理 2, 可知 $v_{t+}-v_t$ 和 v_t-v_{t-} 都是常数的概率为 1. 若注意到 $E(\arctan v_t)=0$, 则知这常数必为 0. 因此,(12.1) 成立. 由此以及定理 1 即知 $\delta(v_t)$ 的不连续点是不存在的.

综合以上所述, 就得分解

$$x_t = u_t + v_t + g(t), \quad g(t) = f(t) + c_t,$$

其中 $g(t)$ 是 t 的函数 (不包含 ω), u_t 和 v_t 是可加过程, $\delta(u_t)$ 是纯粹不连续的, $\delta(v_t)$ 是连续的. 而 $E(\arctan u_t) = E(\arctan v_t) = 0$, 从而得到了 (12.1). 对于 u_t 和 v_t 的关系, 则有下述的定理.

定理 4 两个可加过程 $u_t(t \in T)$ 和 $v_t(t \in T)$ 是独立的 $\Big($即指随机向量 $\prod_t u_t$ 和 $\prod_t v_t$ 是独立的$\Big)$.

证明 由于 $z_{s_i-} - z_{s_{i-1}}, z_{s_i} - z_{s_i-}$ 和 $z_{s_i+} - z_{s_i}$ $(i = 1, 2, \cdots, n)$ 的独立性以及在 §3 里所述有关独立性的性质, 可知 $u_t^{(n)}, t \in T$, 和 $v_t^{(n)}, t \in T$, 是独立的. 故对于任意的 $t_1, t_2, \cdots, t_m$ 来说, m 维的随机向量 $\prod_i u_{t_i}^{(n)}$ 和 $\prod_i v_{t_i}^{(n)}$ 是独立的. 因为本节的引理对于 m 维随机向量也成立, 所以 $\prod_i u_{t_i}$ 和 $\prod_i v_{t_i}$ 是独立的, 进而 $\prod_t u_t$ 和 $\prod_t v_t$ 是独立的.

由上述可知, 要研究可加过程, 只要分别研究 $\delta(x_t)$ 是纯粹不连续的情形和 $\delta(x_t)$ 是连续的情形即可. 前者往往是由下述形状所构成. 固定 $[a, b)$ 中的可数子集 D, 而对 D 的每一点 t, 对应着两个随机变量 ξ_t 和 η_t, 使得所有的 ξ_t 和 $\eta_t (t \in D)$ 都为独立. 并且对 $a \leqslant s < b$ 的 s, 令

$$\xi_s + \sum_{t<s, t \in D} (\xi_t + \eta_t)$$

为收敛型 (§11 末尾的注意). 由此减去 Doob 收敛化常数列, 然后求和, 若令其和为 x_t, 则得具纯粹不连续的 $\delta(x_t)$ 的可加过程. 至于 $\delta(x_t)$ 是连续的场合, 将在以下数节中来说明.

§13 随机过程的可分性[①]

令 $x_t(\omega)(t \in T)$ 为任意的随机过程. 因为对于每个 $t, x_t(\omega)$ 是 ω 的可测函数, 所以

① 关于可分性的详细讨论, 可参见 J. L. Doob, Stochastic Processes, 第 2 章 §2.

—— 校者注

$$A=\{\omega/x_{t_1}(\omega)\leqslant a_1,x_{t_2}(\omega)\leqslant a_2,\cdots,x_{t_n}(\omega)\leqslant a_n\},$$
$$B=\{\omega/\overline{\lim_n}x_{t_n}(\omega)\leqslant a\},$$
$$C=\{\omega/\text{对所有的有理数}t\in T,x_t(\omega)\leqslant 1\},$$

等等, 都是可测集, 故可求出它们的 P 测度. 又若对所有的 $t\in T$,

$$P\{C_t\}=1,\quad C_t=\{\omega/x_t(\omega)\leqslant 1\},$$

则上面的 C 可以写成

$$C=\bigcap_{\substack{t\in T\\ t\text{是有理数}}}C_t,$$

所以它是可列个 P 测度为 1 的集合的交, 因而 $P(C)=1$. 然而, 若考虑以

$$C'=\{\omega/\text{对所有的}t\in T,x_t(\omega)\leqslant 1\}$$

来代替 C 时, 则因 $C'=\bigcap\limits_t C_t$,所以 C 是非可数个 P 测度为 1 的集合的交, 因而无法知道其是否可测, 并可分成三种情形. 有时, 它可以是可测的, 且其 P 测度为 1; 有时它也可以是可测的, 但其 P 测度为 0; 有时, 它也可以是不可测的. 由以上所述, 就可以知道下述结论.

"由随机过程的定义可知, 对于有关至多可数个时点 (t) 的值的事件的概率是可用通常方法加以讨论的, 但是对于有关非可数个时点的值的事件的概率就需要进行特别的讨论."

有关随机过程的重要的情形, 例如

(i) $x_t(\omega)$ 关于 t 是连续的,

(ii) $x_t(\omega)$ 作为 t 的函数是有界的,

(iii) $x_t(\omega)$ 是 t 的增函数,

等等, 都是与非可数个时点有关. 为了讨论这种情形的概率, 就必需对随机过程加以限制. 首先注意到这一点的是 Doob, 他引进了所谓可分性 (separability) 的限制来克服这种困难.

定义 称随机过程 $x_t(t\in T)$ 是可分的, 是指存在 T 的可数子集 S, 使得

(S) $P\{\omega/$对所有的$t\in T,$

$$\varliminf_{\substack{s\to t\\ s\in S}} x_s(\omega) \leqslant x_t(\omega) \leqslant \varlimsup_{\substack{s\to t\\ s\in S}} x_s(\omega)\} = 1.$$

对可分随机过程 $x_t(\omega)(t \in T)$ 来说,(i),(ii) 和 (iii) 中所涉及的事件都是可测的. 显然, 若 (i) 所指的事件的概率等于 1, 则 $x_t, t \in T$, 就是可分的.

称随机过程 $x_t(t \in T)$ 和 $y_t(t \in T)$ **"在弱的意义下相同"** 是指对所有的 $t \in T$, 均有

$$P\{\omega / x_t(\omega) = y_t(\omega)\} = 1.$$

由此, 对任意的 $t_1, t_2, \cdots, t_n \in T$ 和任意的 $E_n \in \boldsymbol{B}^n$, 就有

$$P\{\omega/(x_{t_1}(\omega), \cdots, x_{t_n}(\omega)) \in E_n\} = P\{\omega/(y_{t_1}(\omega), \cdots, y_{t_n}(\omega)) \in E_n\}.$$

更一般地, 对于 $E \in \boldsymbol{B}^T$, 下式也成立

$$P\Big\{\omega / \prod_t x_t \in E\Big\} = P\Big\{\omega / \prod_t y_t \in E\Big\}.$$

然而由于以 T 为定义域的连续函数的全体 C, 虽是 $\mathbb{R}^T$ 的子集, 但不属于 $\boldsymbol{B}^T$, 所以

$$P\Big\{\omega / \prod_t x_t \in C\Big\} = P\Big\{\omega / \prod_t y_t \in C\Big\}$$

不一定成立. 若 $x_t, t \in T$, 不是可分的, 则上式左边的 ω 集合不一定是 P 可测的. 对于 y_t 也是一样的.

在弱的意义下相同的两个随机过程, 即使其中一个是可分的, 另一个也不一定是可分的.

Doob 给出了下述的定理.

定理 1 对任意的随机过程, 存在与它在弱的意义下相同的可分随机过程, 称为原来的随机过程的**可分修正**(separable modification).

因此, 若过程不是可分的, 则可代之以可分修正, 因而只需研究可分过程就够了.

§14 可分 Poisson 过程

如前面所定义那样, 可加过程 $x_t, t \in T$, 当 $x_s - x_t (s > t)$ 的分布是 Poisson 分布 $P(\lambda(s-t)$ 时, 叫做 **Poisson 过程**(Poisson process). 若它是

可分的, 则叫做**可分 Poisson 过程**. 因为 Poisson 过程存在, 所以取其可分修正, 就知道可分 Poisson 过程也存在.

定理 1 若令 $x_t, t \in T = [a, b)$, 为可分 Poisson 过程, 则其样本过程具概率为 1 地是跃度 1 的增的阶梯函数 (但是跳跃点上的值处于左右两极限之间).

证明 由可分性的定义, 可知存在 T 的可数子集 S, 使下述的 ω 集合的 P 测度为 1:

$$\Omega' = \left\{\omega / \varliminf_{\substack{s \to t \\ s \in S}} x_s(\omega) \leqslant x_t(\omega) \leqslant \varlimsup_{\substack{s \to t \\ s \in S}} x_s(\omega), t \in T\right\}.$$

当 t 为固定时,

$$\Omega_t = \{\omega / x_t(\omega) = \text{非负整数}\}$$

的 P 测度就为 1. 这是由于 x_t 的分布是 $P(\lambda(t-a))$ 的缘故. 又因 $x_t - x_u (t > u)$ 的分布为 $P(\lambda(t-s))$, 所以

$$\Omega_{ut} = \{\omega / x_t(\omega) - x_u(\omega) \geqslant 0\}$$

的 P 测度也应为 1. 而 S 是可数集合, 所以

$$\Omega'' = \bigcap_{s \in S} \Omega_s \cap \bigcap_{\substack{s<t \\ s,t \in S}} \Omega_{st} \cap \Omega'$$

的 P 测度也为 1. 若 $\omega \in \Omega''$, 则作为 t 的函数, $x_t(\omega)$ 就是跃度为 $1, 2, 3, \cdots$ 的增的阶梯函数.

若能证明使得 $x_t(\omega)$ 的跃度大于或等于 2 的 ω 集合 N 的 P 测度为 0, 那么证明就告完毕. 这只需证明 $x_t(\omega)$ 在 $a \leqslant t \leqslant t_0$ 内具有跃度大于等于 2 的跳跃的 ω 集合 $N(t_0)$ 的 P 测度为 0 即可. 现在若设

$$E_{nk} = \left\{\omega \middle/ x\left(a + \frac{k}{n}(t_0 - a), \omega\right) - x\left(a + \frac{k-1}{n}(t_0 - a), \omega\right) \geqslant 2\right\}, \quad 1 \leqslant k \leqslant n,$$

则得

$$N(t_0) \subset \bigcup_k E_{nk},$$

故得

$$P(N(t_0)) \leqslant \sum_{k=1}^{n} P(E_{nk}) = O\left(n\left(\frac{1}{n}\right)^2\right) \to 0 \quad (n \to \infty).$$

定义 令 $x_t, t \in T$, 为任意的随机过程, 且对任意的 $\varepsilon > 0$,

$$P\{\omega / |x_t - x_s| > \varepsilon\} \to 0 \quad (t \to s)$$

时, 就称随机过程在 s 上**依概率连续**(continuous in probability). 若在 T 的所有的点上依概率连续时, 则称 $x_t, t \in T$, 为**依概率连续的随机过程**.

若对上述的 (可分)Poisson 过程来说, 由于

$$P\{\omega / |x_t - x_s| > \varepsilon\} = O(|t - s|) \to 0 \quad (t \to s),$$

所以它是依概率连续的. 但是可分 Poisson 过程的样本过程恰恰不是连续函数而是阶梯函数. 粗略地说, 就是每个样本过程都含有不连续点, 但因不连续点随 ω 而变动, 所以若着眼于一个点, 则在这点上为不连续的概率为 0. 实际上, 若令 $x_{t+0} = x_{t-0}$ 的 ω 集合为 N_t 时, 虽然 $P(N_t) = 1$, 但 $P(\bigcap_t N_t) = 0$.

若 x_t 是依概率连续的可加过程, 并且 $x_s - x_t (s > t)$ 的分布为 Poisson 分布, 则 x_t 叫做**广义 Poisson 过程**. 若令 $\lambda(t) = E(x(t))$, 则由过程的依概率连续性, $\lambda(t)$ 就成为 t 的连续函数. $x_s - x_t (s > t)$ 的分布就是 $P(\cdot; \lambda(s) - \lambda(t))$. 前述的 Poisson 过程就是 $\lambda(t) = \lambda(t - a)$ 这一特殊场合. 此时, 称它为**对时间齐次**(temporally homogeneous) 的 Poisson 过程.

定理 1′ 定理 1 相同的结论, 对广义可分 Poisson 过程来说也是成立的.

反之, 有下述的定理.

定理 2 若有一依概率连续的可加过程, 其样本过程具概率为 1 地是跃度为 1 的增的阶梯函数, 则它就是广义可分 Poisson 过程.

证明 令 $x_t, t \in T = [a, b)$, 为所讨论的可加过程. 由于假定样本过程是阶梯函数, 可分性是明显的, 因此, 只要证明 $x_s - x_t (s > t)$ 的分布是 Poisson 分布就行了. 由依概率连续性, 知对 $u \in T$ 和 $\varepsilon > 0$, 存在 $\delta = \delta(\varepsilon, u)$, 并且当 $|v - u| < \delta$ 时, 使得

$$P\{\omega / |x_v - x_u| > \varepsilon\} < \varepsilon,$$

但由 Borel 覆盖定理, 可知在 $t \leqslant u \leqslant s$ 内可以取得 $\delta(\varepsilon, u)$ 与 u 无关. 把区间 $[t, s]$ 划分为 n 等分, 令 x_t 在各小区间上的增量为 $y_{n1}, y_{n2}, \cdots, y_{nn}$, 且设 $y = y_{n1} + \cdots + y_{nn} = x_s - x_t$. 令 y'_{nk} 随着 $y_{nk} \leqslant 1$ 或者 $\geqslant 2$ 而等于 y_{nk} 或者 0, 又设

$$y'_n = y'_{n1} + y'_{n2} + \cdots + y'_{nn}.$$

根据有关 x_t 的样本过程的假定, 而得

$$P(y'_n \to y) = 1.$$

其次, 若设 $p_{nk} = P\{\omega / y'_{nk} = 1\}$, 则 $p_{nk} = P\{\omega / y_{nk} = 1\}$, 所以由上述而知, 极限 $p_{nk} \to 0 (n \to \infty)$ 对 k 一致地成立. 又因 $\{y_{nk}, k = 1, 2, \cdots, n\}$ 是独立的, 所以 $\{y'_{nk}, k = 1, 2, \cdots, n\}$ 也是独立的.

(i) $\sum_k p_{nk} \to \lambda$(有限) 时,

$$\begin{aligned} E(\mathrm{e}^{izy}) &= \lim_n E(\mathrm{e}^{izy'_n}) = \lim_n \prod_k E(\mathrm{e}^{izy'_{nk}}) \\ &= \lim_n \prod_k (1 - p_{nk} + \mathrm{e}^{iz} p_{nk}) = \lim_n \prod_k (1 + p_{nk}(\mathrm{e}^{iz} - 1)). \end{aligned}$$

由 $\sum_k p_{nk} \to \lambda$ 和 $\sum_k p^2 \leqslant \max_k p_{nk} \sum_k p_{nk} \to 0$, 可知上面的最后一式变为 $\exp\{\lambda(\mathrm{e}^{iz} - 1)\}$, 于是知道 y 的分布是 Poisson 分布.

(ii) $\sum_k p_{nk}$ $(n = 1, 2, \cdots)$ 的子序列有有限的极限时, 也与 (i) 相同.

(iii) $\sum_k p_{nk} \to \infty$ 时. 因为 p_{nk} 当 $n \to \infty$ 时关于 k 一致地变小, 所以对任意的 $\lambda > 0$ 和 n, 可以定出 $K = K(n, \lambda)$, 使得

$$\sum_{k=1}^{K} p_{nk} \to \lambda.$$

若设 $y''_n = \sum_{k=1}^{K} p_{nk}$, 则与 (i) 同样地可得

$$\begin{aligned} |E(\mathrm{e}^{izy})| &\leqslant \lim_n \Big| \prod_{k=1}^{K} E(\mathrm{e}^{izy'_{nk}}) \Big| = |\exp\{\lambda(\mathrm{e}^{iz} - 1)\}| \\ &= \exp\{\lambda(\cos z - 1)\}. \end{aligned}$$

因为 λ 可任意取无论多大的数, 若令 $\lambda \to \infty$ 时, 则在 $0 < z < 2\pi$ 的范围内右边为 0, 所以 $|E(e^{izy})| = 0$. 若令 $z \downarrow 0$, 则 1=0, 于是导致了矛盾的结果. 故不可能发生 (iii) 的情形.

§15 可分 Wiener 过程

关于可分 Wiener 过程的定义和其存在性, 已无需加以说明了.

定理 1 可分 Wiener 过程的样本过程具概率为 1 地为连续.

证明 令 $x_t, t \in T = [a, b)$, 为 Wiener 过程, 且令在定义可分性时所用的可数集合为 S. 也就是说

$$\Omega' = \left\{\omega / \varliminf_{\substack{s\to t\\ s\in S}} x_s(\omega) \leqslant x_t(\omega) \leqslant \varlimsup_{\substack{s\to t\\ s\in S}} x_s(\omega), t \in T\right\}$$

的概率是 1. 由正态分布的性质得

$$P\{\omega/|x_\beta - x_\alpha| > \varepsilon\} = o(\beta - \alpha),$$

对于 $\alpha = \alpha_0 < \alpha_1 \cdots < \alpha_n = \beta$, 因为 $y_\nu = x_{\alpha_\nu} - x_{\alpha_{\nu-1}}, \nu = 1, 2, \cdots, n$, 是独立的, 所以由 Ottaviani 定理得

$$P\{\omega/\max_\nu |x_{\alpha_\nu} - x_\alpha| > 2\varepsilon\} \leqslant 2P\{\omega/|x(\beta) - x(\alpha)| > \varepsilon\}.$$

其次, 若在 $[\alpha, \beta]$ 内有可数个点 $t_1, t_2, \cdots$ 时, 则可按照大小的顺序把 t_1, $t_2, \cdots, t_n$ 重新排列, 然后利用上述的结果, 再令 $n \to \infty$ 便得

$$P\{\omega/\sup_\nu |x_{t_\nu} - x_\alpha| > 2\varepsilon\} \leqslant 2P\{\omega/|x_\beta - x_\alpha| > \varepsilon\}.$$

因此

$$P\left\{\omega/ \sup_{t\in[\alpha,\beta]\bigcap S} |x(t) - x(\alpha)| > 2\varepsilon\right\} = o(\beta - \alpha),$$

由可分性得

$$P\left\{\omega/ \sup_{\alpha\leqslant t\leqslant\beta} |x(t) - x(\alpha)| > 2\varepsilon\right\} = o(\beta - \alpha).$$

那么, 若 $b < \infty$, 则把 $[a, b)$ 等分为 n 个小区间 $I_{n\nu} = [a_{n,\nu-1}, a_{n,\nu}], \nu = 1, 2, \cdots, n$, 就得

$$P\Big\{\omega/\max_{\nu=1}^{n}\sup_{t\in I_{n\nu}}|x_t - x_{\alpha_{n,\nu-1}}| > 2\varepsilon\Big\}$$
$$\leqslant\sum_{\nu}P\Big\{\omega/\sup_{t\in I_{n\nu}}|x_t - x_{\alpha_{n,\nu-1}}| > 2\varepsilon\Big\} = n\cdot o\Big(\frac{1}{n}\Big) = o(1).$$

若 $|t-s| < (b-a)/n$, 则 t 和 s 落入同一的 $I_{n\nu}$ 中或者分别落入邻近的 $I_{n\nu}$ 中, 所以

$$P\Big\{\omega/\sup_{|t-s|<(b-a)/n}|x_t - x_s| > 4\varepsilon\Big\} = o(1).$$

因为这个 ω 集合随 n 而减少, 所以

$$P\Big\{\omega/\lim_{n}\sup_{|t-s|<(b-a)/n}|x(t) - x(s) > 4\varepsilon\Big\} = 0.$$

令 $\varepsilon\to 0$ 就得定理的结论.

当 $b=\infty$ 时, 若设

$$\Omega_k' = \{\omega/x_t(\omega)\text{连续于}a\leqslant t\leqslant k\},$$
$$\Omega' = \{\omega/x_t(\omega)\text{连续于}a\leqslant t<\infty\},$$

则 $\Omega_k'\downarrow\Omega'$. 由以上所述就得 $P(\Omega_k')=1$, 故 $P(\Omega')=1$.

与 Poisson 过程的场合一样, 可以定义广义 Wiener 过程为依概率连续的可加过程, 而且 $x_s - x_t (s>t)$ 的分布为正态分布. 若令 x_t 即 $x_t - x_a$ 的分布为 $N(\cdot;m(t),v(t))$, 即 $x_s - x_t$ 的分布为 $N(\cdot;m(s)-m(t),v(s)-v(t))$. 由依概率连续性的假定可推出 $m(t)$ 和 $v(t)$ 是连续的. 当然 $v(t)$ 是增函数. 关于广义 Wiener 过程及广义可分 Wiener 过程的存在, 恐怕不需要重新说明了.

定理 1′ 定理 1 对于广义 Wiener 过程来说也成立.

反之, 有下述的定理.

定理 2 若有一依概率连续的可加过程, 其样本过程具概率为 1 地连续, 则它是广义可分 Wiener 过程.

证明 令 $x_t, t\in T=[a,b)$, 为所讨论的可加过程. 由于假定样本过程是连续的, 因而可分性是明显的, 只要证明 $x_s - x_t (s>t)$ 的分布是正态分布就行了. 考虑到连续函数在有界区间上的一致连续性, 因而对任意的 $\varepsilon>0$, 可以找到 $\delta=\delta(\varepsilon)$, 使得

$$P\{\omega/u,v\in[t,s],|u-v|<\delta\Rightarrow|x(u)-x(v)|<\varepsilon\}>1-\varepsilon^{①}.$$

于是就可取一列 $\varepsilon_1>\varepsilon_2>\cdots\to 0$, 并可取正整数 $p(n)$ 使 $(s-t)/p(n)=o(\varepsilon_n)$. 若把 $[t,s)$ 等分为 $p(n)$, 而令 $t=t_{n0}<t_{n1}<\cdots<t_{np(n)}=s$, 然后依 $y_{n\nu}=x_{t_{n\nu}}-x_{t_{n,\nu-1}}$, 或 $|y_{n\nu}|<\varepsilon_n$, 而定义 $y'_{n\nu}=y_{n\nu}$ 或者等于 0, 再设 $y'_n=\sum\limits_{\nu}y'_{n\nu}$, 则得

$$P\{\omega/y\neq y'_n\}<\varepsilon_n,$$

所以考虑到 $\{y'_{n\nu}\}$ 的独立性, 就得

$$E(\mathrm{e}^{izy})=\lim_n E(\mathrm{e}^{izy'_n})=\lim_n\prod_\nu E(\mathrm{e}^{izy'_{n\nu}}).\qquad(*)$$

(i) 设 $m_{n\nu}=E(y'_{n\nu}),v_{n\nu}=V(y'_{n\nu}),m_n=\sum\limits_\nu m_{n\nu},v_n=\sum\limits_\nu v_{n\nu}$, 若 $m_n\to m$(有限), $v_n\to v$(有限), 则由 $(*)$ 得

$$\begin{aligned}E(\mathrm{e}^{izy})&=\lim_n \mathrm{e}^{izm_n}\prod_\nu E(\mathrm{e}^{iz(y'_{n\nu}-m_{n\nu})})\\&=\mathrm{e}^{izm}\lim_n\prod_\nu\left(1-\frac{v_{n\nu}}{2}z^2+v_{n\nu}\cdot O(\varepsilon_n)\right)\\&=\mathrm{e}^{izm}\lim_n\prod_\nu \mathrm{e}^{-\frac{v_{n\nu}}{2}z^2+v_{n\nu}O(\varepsilon_n)}\\&=\mathrm{e}^{izm}\lim_n \mathrm{e}^{-\frac{v_n}{2}z^2+v_nO(\varepsilon_n)}\\&=\mathrm{e}^{izm-\frac{v}{2}z^2},\end{aligned}$$

故 y 的分布是正态分布.

(ii) m_n 和 v_n 的子序列分别接近于有限值时, 也与 (i) 相同.

(iii) 当 v_n 或者其子序列接近于有限值 v 时, 与 (i) 的做法一样, 就可使得 y'_n-m_n 或者它的子序列的分布接近于 $N(0,v)$, 然而因 y'_n 的分布接近于 y 的分布, 所以 m_n 或者它的子序列接近于某个有限值 m, 于是就归结到 (ii) 的场合.

(iv) 若 $v_n\to\infty$, 由于 $|v_{n\nu}|\leqslant 4\varepsilon_n^2$, 所以对任意的 v 可以定出 $q(n)$ 使 $\sum\limits_{\nu=1}^{q(n)}v_{n\nu}\to v$. 由 $(*)$ 得

① 这个不等式可仿照实变函数论中的叶果洛夫定理的证明方法由过程的依概率连续性即可推出. —— 校者注

$$|E(\mathrm{e}^{izy})| \leqslant \overline{\lim_n} \prod_{\nu=1}^{q(n)} |E(\mathrm{e}^{izy'_{n\nu}})| = \mathrm{e}^{-\frac{v}{2}z^2}.$$

令 $v \to \infty$, 就得 $|E(\mathrm{e}^{izy})| = 0$. 这是矛盾的结果. 于是 (iv) 的情形不会发生.

§16 依概率连续的可加过程和无穷可分分布律

早已在 §12 中证明了这样的分解: 一般的可加过程 $x_t, t \in T$, 可以分解为

$$x_t = u_t + v_t + g(t),$$

而 $\delta(u_t)$ 为连续, 且 $E(\arctan u_t) = 0; \delta(v_t)$ 为纯粹不连续, 且 $E(\arctan v_t) = 0; g(t)$ 只是 t 的函数, 而且 $u_t(t \in T)$ 与 $v_t(t \in T)$ 是独立的. 因为 v_t 的构造已在 §12 中说明了, 所以就剩下考察 u_t 的问题. 由前可知对 u_t 有

$$P\{\omega/u_{t-} = u_t = u_{t+}\} = 1, \quad t \in T. \tag{16.1}$$

因此它是依概率连续的. 但是, 正如对 Poisson 过程所指出的情形一样, (16.1) 并不能表示样本过程的连续性.

依概率连续的可加过程究竟是怎样的过程呢? 下面就来研究这个问题.

定理 1 令 x_t 为依概率连续的可加过程. 对于任意的 t, 取序列 $t_1 < t_2 < \cdots \to t$, 则 x_{t_n} 几乎处处收敛于 x_t. 对于 $t_1 > t_2 > \cdots$ 的序列来说也是一样的. 当然, 几乎处处收敛中的例外 ω 集合, 一般是与 t 以及序列的选择有关的.

证明 由依概率连续的假定, 可知x_{tn}依概率收敛于x_t. 对于序列$t_1 < t_2 < \cdots \to t$, 因 $x_{t_n} = x_{t_1} + (x_{t_2} - x_{t_1}) + \cdots + (x_{t_n} - x_{t_{n-1}})$ 中的各项是独立的, 故由 §10 的定理 3, 可知 x_{t_n} 是几乎处处收敛的. 因而 x_{t_n} 几乎处处收敛于 x_t. 对于 $t_1 > t_2 > \cdots$ 的情形也是一样.

我们已经知道, 广义 Poisson 过程和广义 Wiener 过程都是依概率连续, 而且其增量 $x_s - x_t(s > t)$ 的分布分别是 Poisson 分布和正态分布. 因此, 由于考虑一般的依概率连续的可加过程的增量 $x_s - x_t$ 的分布, 就可

设想它是包含 Poisson 分布以及正态分布在内的一般的分布. 这种分布叫做无穷可分分布.

若一维分布 Φ 满足

$$\int_{|\xi|>\varepsilon} \Phi(\mathrm{d}\xi) < \varepsilon,$$

则记之为 $\Phi \in U(\varepsilon)$. 因此, Φ_n 趋向于单位分布, 就等价于: 对任意的 $\varepsilon > 0$, 可以找到 $n_0(\varepsilon)$, 当 $n > n_0(\varepsilon)$ 时, $\Phi_n \in U(\varepsilon)$.

将依概率连续的可加过程的增量 $x_s - x_t$ 的分布记为 Φ_{ts}. 根据依概率连续性, 对任意的 $\varepsilon > 0$ 和 u, 可以找到 $\delta = \delta(\varepsilon, u)$, 使得

$$|v-u| < \delta \Rightarrow P\{\omega/|x_v - x_u| > \varepsilon\} < \varepsilon,$$

但由 Borel 覆盖定理可知, 对 $t \leqslant u \leqslant s$ 而言, 可以取 $\delta(\varepsilon, u)$ 为与 u 无关的 $(\delta(\varepsilon))$. 若把区间 $[t, s)$ 分得充分地小, 使 $t = u_0 < u_1 < \cdots < u_n = s$, 且 $u_i - u_{i-1} < \delta(\varepsilon)$, 则由上述可得

$$\Phi_{u_{i-1}u_i} \in U(\varepsilon), \quad i = 1, 2, \cdots, n.$$

又由可加性, 得

$$\Phi_{ts} = \Phi_{u_0u_1} * \Phi_{u_1u_2} * \cdots * \Phi_{u_{n-1}u_n}.$$

粗略地说, 就是 Φ_{ts} 可表为充分接近单位分布的分布的卷积.

现在脱离可加过程来考虑. 若有分布 Φ, 对任意的 $\varepsilon > 0$, 成立着形状如

$$\Phi = \Phi_1 * \Phi_2 * \cdots * \Phi_n, \qquad \Phi_1, \Phi_2, \cdots, \Phi_n \in U(\varepsilon)$$

的分解式时, Φ 就叫做**无穷可分的**(infinitely divisible). 由

$$N(\cdot; m, v) = N_1 * \cdots * N_n, \quad N_1 = N_2 = \cdots = N_n = N\left(\cdot; \frac{m}{n}, \frac{v}{n}\right)$$

$$P(\cdot; \lambda) = P_1 * \cdots * P_n, \quad P_1 = P_2 = \cdots = P_n = P\left(\cdot; \frac{\lambda}{n}\right)$$

便知正态分布和 Poisson 分布都是无穷可分的. 依概率连续的可加过程的增量的分布 (上述的 Φ_{ts}) 也显然是无穷可分的. 反之, 有下述定理.

定理 2 对于无穷可分的分布 Φ, 可以定出依概率连续的可加过程 $x_t, t \in [0, 1]$, 使得 $x_1 (= x_1 - x_0)$ 的分布等于 Φ.

证明 当 $\Phi = \delta(\cdot; m)$ 时, 设 $x_t(\omega) \equiv m \cdot t$ 就行了. 因此不考虑这种特殊情形, 而假定 $\delta(\Phi) > 0$. 由

$$\int \arctan(\xi - c)\Phi(\mathrm{d}\xi) = 0$$

唯一定出了一个常数 c(这在讨论 Doob 收敛常数列时已用过), c 叫做 Φ 的**中位数**$\alpha(\Phi)$. 当 $\alpha(\Phi) = c$ 时, $\Phi * \delta(\cdot; m)$ 的中位数就为 $c + m$. 为了证明此定理, 不妨假定 $\alpha(\Phi) = 0$. 这是因为对一般的 Φ 来说, 若考虑 $\Phi * \delta(\cdot; -c), c = \alpha(\Phi)$ 时, 则也是无穷可分分布, 而且其中位数为 0, 所以若令相应的可加过程为 x_t, 则 $y_t = x_t + ct$ 即为对应于 Φ 的可加过程.

令 [0,1] 中的全体有理数为 S, 规定分布序列 $\Phi = \{\Phi_{ts}, t \leqslant s, t, s \in S\}$ 满足下述四个条件:

(i) $\Phi_{01} = \Phi$,

(ii) $s \leqslant t \leqslant u \Rightarrow \Phi_{st} * \Phi_{tu} = \Phi_{su}$,

(iii) $\alpha(\Phi_{0t}) = 0$,

(iv) $\delta(\Phi_{0t}) = t\delta(\Phi)$.

取 $\varepsilon_n \to 0$ 的正数列 $\{\varepsilon_n\}$, 再令 Φ 对 ε_n 的分解为

$$\Phi = \Phi_{n1} * \Phi_{n2} * \cdots * \Phi_{np(n)}, \quad \Phi_{ni} \in U(\varepsilon_n).$$

若有必要, 就以 $\Phi_{ni} * \delta(\cdot; m_i)$ 代替 Φ_{ni}, 因而可以假定

$$\alpha(\Psi_{ni}) = 0, \quad \Psi_{ni} = \Phi_{n1} * \cdots * \Phi_{ni}, \quad i = 1, 2, \cdots, n.$$

又可以假定所有的 Φ_{ni} 都不是 δ 分布, 因为 $\delta(\Phi_{ni})$ 随 i 而真正地增加. 对 t 确定 i, 使得

$$\delta(\Psi_{n,i-1})/\delta(\Phi) \leqslant t < \delta(\Psi_{n,i})/\delta(\Phi),$$

同样, 对 $s(> t)$ 也可确定满足上式的 j. 令

$$\Phi_{ts}^{(n)} = \Phi_{n,i} * \cdots * \Phi_{n,j-1}(i < j), \quad \Phi_{ts}^{(n)} = \text{单位分布}\ (i = j).$$

$\Phi^{(n)} = \{\Phi_{ts}^{(n)}, t, s \in S, t \leqslant s\}$ 可以看做是 Φ 的近似系列. 对于 $\Phi^{(n)}$, 虽然 (i)、(ii) 和 (iii) 也成立, 但 (iv) 就不过是近似的而已. 从 $\Phi^{(n)}$ 取出子序列而定义 Φ 为其极限, 这就是欲求的目的. 为此先引进下述的引理.

引理　若 $\Phi_n * \Psi_n = \Phi, n = 1, 2, \cdots$, 则存在 $\{\Phi_n\}$ 的子序列 $\{\Phi_{n'}\}$ 和实数列 $\{m(n)\}$, 使得 $\{\Phi_{n'} * \delta(\cdot, m(n))\}_n$ 收敛于某个 Φ_∞. 特别当 $\alpha(\Phi_n) = 0$ 时, 可以假设 $m(n) = 0$.

证明　令 Φ_n 的分布函数为 $F_n(x)$. 因为由 $\Phi_n * \Psi_n = \Phi$ 得出 $[\Phi_n * \delta(\cdot; m)] * [\Psi_n * \delta(\cdot; -m)] = \Phi$, 所以假定

$$F_n(-0) \leqslant 1/2, \quad F_n(0) \geqslant 1/2$$

也不失普遍性. 由这条件得出

$$F_n(y) - F_n(x) > 1/2 \Rightarrow x \leqslant 0 \leqslant y. \tag{16.2}$$

由假定 $\Phi_n * \Psi_n = \Phi$, 便得

$$\int_{-\infty}^{\infty} [F_n(l - x) - F_n(-l - x)] \Psi_n(\mathrm{d}x) = \Phi[-l, l].$$

取 l 足够大, 使右边大于 1/2, 然后适当地取 $m(n)$ 就得

$$F_n(l - m(n)) - F_n(-l - m(n) - 0) \geqslant \Phi[-l, l] \geqslant 1/2.$$

故由 (16.2) 便得 $-l \leqslant m(n) \leqslant l$. 因此, 可适当地取 $\{n\}$ 的子序列 $\{n'\}$ 使 $\{m(n')\}_n$ 收敛于某个 m, 另外可取 $\{n'\}$ 的子序列 (再次记为 $\{n'\}$) 使 $\{F_{n'}(x - m(n_1'))\}_n$ 除在可数例外点以外, 收敛于单调增的右连续函数 $F_0(x)$. 因为 $m(n') \to m, \{F_n'(x - m)\}$ 也在上述的意义下趋近于 $F_0(x)$, 因而 $\{F_n'(x)\}$ 除在可数例外点以外, 趋近于 $F(x) \equiv F_0(x + m)$. 又由 $0 \leqslant F(x) \leqslant 1$ 可得

$$F(\infty) - F(-\infty) \geqslant F_0(l) - F_0(-l - 0) \geqslant \Phi[-l, l].$$

其次, 令 L 为大于 l 的任意的数, 而对 L 和 $\{F_n'(x)\}_n$ 重复上述的论证, 就可得到 $\{F_n'(x)\}$ 的子序列 $\{F_n''(x)\}$, 使其除在可数例外点以外, 趋近于某个 $G(x)$, 而且

$$G(\infty) - G(-\infty) \geqslant \Phi[-L, L].$$

然因 $\{F_n'(x)\}$ 本来就趋近于 $F(x)$, 所以 $G = F$, 故得

$$F(\infty) - F(-\infty) \geqslant \Phi[-L, L].$$

因为 L 是任意的, 所以令 $L \to \infty$ 便得 $F(\infty) - F(-\infty) = 1$. 若令 F 所对应的一维分布为 Φ_∞, 则 $\{\Phi_{n'}\}_n$ 就趋近于 Φ_∞. 若 $\alpha(\Phi_n) = 0$, 则可由 $\Phi_{n'} * \delta(\cdot; m(n)) \to \Phi_\infty$ 得出 $m(n) \to \alpha(\Phi_\infty)$, 所以 $\{\Phi_{n'}\}$ 本身就成为收敛序列.

现在回到定理的证明上去. 因为 $\Phi_{st}^{(n)}*[\Phi_{0t}^{(n)}*\Phi_{t1}^{(n)}] = \Phi$, 所以由引理就知, 可取 $\{n\}$ 的子序列 $\{n'\}$ 和数列 $\{m_n(s,t)\}$ 使得 $\{\Phi_{st}^{(n')}*\delta(\cdot; m_n(s,t))\}_n$ 收敛. 此子序列 $\{n'\}$ 的选择一般是与 $s, t(\in S)$ 有关的, 但考虑到 S 是可数的, 而且应用对角线的方法, 就可以取得所有的 s, t 共同适用的子序列. 由引理的后一半, 就可得出 $m_n(0,t) = 0$. 现在证明 $m_n(s,t) = 0$. 为此, 只要证明 $\{m_n(s,t)\}$ 收敛即可. 显然

$$\Phi_{0s}^{(n')} * [\Phi_{st}^{(n')} * \delta(\cdot; m_n(s,t))] = \Phi_{0t}^{(n')} * \delta(\cdot; m_n(s,t)),$$

而由于可选择 $\{n'\}$ 左边及右边的 $\Phi_{0t}^{(n')}$ 收敛, 故 $\{m_n(s,t)\}_n$ 也非收敛不可. 因而 $\{\Phi_{st}^{(n')}\}_n$ 也收敛. 若记其极限分布为 Φ_{st}, 则当 $s \leqslant t \leqslant u$ 时, 由 $\Phi_{st}^{(n)} * \Phi_{tu}^{(n)} = \Phi_{su}^{(n)}$ 可得出 $\Phi_{st} * \Phi_{tu} = \Phi_{su}$. 因为 $\alpha(\Phi_{0t}^{(n)}) = 0$, 所以 $\alpha(\Phi_{0t}) = 0$. 又因

$$\delta(\Phi_{0t}^{(n)}) \leqslant t \cdot \delta(\Phi) < \delta(\Phi_{0t}^{(n)} * \Phi_{n,i(n,t)}),$$

此处 $\Phi_{n,i(n,t)} \in U(\varepsilon_n)$, 所以 $\{\Phi_{n,i(n,t)}\}$ 趋近于单位分布. 令上面不等式中的 $n \to \infty$, 就得

$$\delta(\Phi_{0t}) \leqslant t \cdot \delta(\Phi) \leqslant \delta(\Phi_{0t}), \quad 即\ \delta(\Phi_{0t}) = t \cdot \delta(\Phi).$$

因而 $\Phi = \{\Phi_{s,t}; s < t, s, t \in S\}$ 满足 (i),(ii),(iii) 和 (iv).

由 §6 定理 2 便知, 存在可加过程 $y_t, t \in S$, 而且 $y_t - y_s (t > s)$ 的分布为 Φ_{st}. (在 §6 的定理中, t 的区域 T 是区间, 但其证明仍然适用于 $T = S$ 的场合.) 其次, 对任意的 $t \in [0,1]$ 取 $t_n \in S$ 且使 $t_n \uparrow t$, 则 $\{y_{t_n}\}$ 是可加序列, 又由 $\alpha(y_{t_n}) = 0$ 及 $\delta(y_{t_n}) = t_n\delta(\Phi) \leqslant t\delta(\Phi)$, 可知 $\{y_{t_n}\}$ 为几乎处处收敛. 若令其极限为 x_t, 则 $x_t, 0 \leqslant t \leqslant 1$, 是可加过程, 又因 $\alpha(x_t) = 0$ 和 $\delta(x_t) = t\delta(\Phi)$(关于 t 连续), 所以 x_t 依概率连续.

§17 依概率连续的可分可加过程的构造

如早在 §13 中讨论过的那样, 要研究可加过程, 只要研究可分可加过

程就够了. 本节的目的是详细地考察依概率连续的可分可加过程的构造. 由于证明过于复杂, 所以从略.

定理 1[①] 若令 $x_t, t \in [a,b)$, 为依概率连续的可分可加过程, 则其样本过程具概率为 1 地为第一种不连续函数.

若对所有的 $t, f(t-0)$ 和 $f(t+0)$ 存在, 而且 $f(t)$ 处于这两极限之间, 则 $f(t)$ 叫做**第一种不连续函数**. 在 (t,u) 平面上画出

$$u = f(t+0) - f(t-0), \quad a \leqslant t \leqslant b$$

的图像. 所谓图像, 一般地说, 也就是不连续点的集合, 它是带状区域 D : $a \leqslant t \leqslant b, -\infty < u < \infty$, 的一个子集 $G(f)$. 对 D 的任意 Borel 子集 E, 记 $E \cap G(f)$ 中的点的个数为 $N_f(E)$. 若 $E \cap G(f)$ 是无限集合, 则不管它的势是怎样, 总设 $N_f(E) = \infty$. $N_f(E)$ 就是图像在 E 中的点的个数. $N_f(E)$ 可以视为 D 上的测度. 特别是 E**真正离开**t 轴 (即距离为正) 时, 由第一种不连续函数的定义, 立即可以证明 $N_f(E)$ 是有限的. 其次, 以 $S_f(E)$ 来表示 $E \cap G(f)$ 中的点的 u 坐标的和数. 也就是

$$\begin{aligned} S_f(E) &= \sum_{(t,u) \in E \bigcap G(f)} u = \sum_{(t, f(t+0) - f(t-0)) \in E} (f(t+0) - f(t-0)) \\ &= \int_E u N_f(\mathrm{d}t\mathrm{d}u). \end{aligned}$$

当 E 处在上半平面或者下半平面时, $S_f(E)$ 可以取 $+\infty$ 或 $-\infty$; 当 E 跨及 t 轴的上下时, 有时就不能确定. 但是, 若 E 真正离开 t 轴时, 则 $S_f(E)$ 可以确定, 而且其值是有限的.

若 $x_t(\omega), t \in [a,b)$, 是依概率连续的可分可加过程, 则其样本过程具概率为 1 地为第一种不连续函数, 所以对此可以定义上述的 $N_x(E)$ 和 $S_x(E)$. 这时它们都与 ω 有关, 且是 ω 的可测函数.

定理 2 $N_x(E)$ 的分布是 Poisson 分布. 但此时规定恒等于 ∞ 的随机变量的分布为 $P(\cdot;\infty)$.

定理 3 若 $E_1, E_2, \cdots$ 互不相交, 则 $N_x(E_1), N_x(E_2), \cdots$ 相互独立, 又若设 $E = \sum\limits_n E_n$, 则

① 参见 J. L. Doob, Stochastic Processes,1953, 第 8 章定理 7.2. —— 校者注

$$N_x(E) = \sum_n N_x(E_n).$$

以 N 表示随机变量的系 $\{N_x(E)\}_E$. 由上述二定理可知, N 非常类似于 Poisson 过程. 在这种意义下, N 叫做 Poisson 可加系. 若令 $N_x(E)$ 的均值为 $n(E)$, 则 $N_x(E)$ 的分布为 $P(\cdot; n(E))$. 由 $N_x(E)$ 的可加性便知, $n(E)$ 是 $D: a \leqslant t \leqslant b, -\infty < u < \infty$, 上的测度. 因为当 E 真正离开 t 轴时 $N_x(E)$ 恒为有限, 所以 $n(E)$ 也是有限的. $n(E)$ 在 E 不离开 t 轴时有可能为 ∞, 而其所大到的程度由下述定理给出.

定理 4 $\displaystyle\int_{u=-1}^{1}\int_{t=a}^{b} u^2 n(\mathrm{d}t\mathrm{d}u) < \infty.$

令从 a 到 t 的 x_t 的跳跃点中跃度绝对值在 $1/n$ 以上的总和为 $S_n(t)$, 即

$$S_n(t) = \int_{|u|>1/n}\int_{\tau=a}^{t} uN(\mathrm{d}\tau\mathrm{d}u).$$

这就是可加过程. 当 $n \to \infty$ 时, 一般地说 $S_n(t)$ 不收敛, 但适当地加减一个 t 的函数 (与 ω 无关), 就可以使之收敛. 这就是下述定理.

定理 5 若设

$$S_n^*(t) = S_n(t) - \int_{|u|>1/n}\int_{\tau=a}^{t} \frac{u}{1+u^2} n(\mathrm{d}\tau\mathrm{d}u),$$

则当 $n \to \infty$ 时, $S_n^*(t)$ 具概率为 1 地关于 t 一致收敛. 若令这极限过程为 $S(t)$, 则这也是可加过程, 而且 $u = S(t+0) - S(t-0)$ 的图像和 $u = x_{t+0} - x_{t-o}$ 的图像具概率为 1 地是一致的.

由此定理可知, $y_t = x_{t+0} - S(t+0)$ 具概率为 1 地为 t 的连续函数.

定理 6 y_t 是可分正态随机过程, 且与上述的 Poisson 可加系 $N = \{N_x(E)\}_E$ 是独立的.

综合以上所述就得

$$x_{t+0} = y_t + \lim_n \int_{|u|>1/n}\int_{\tau=a}^{t} \left[uN(\mathrm{d}\tau\mathrm{d}u) - \frac{u}{1+u^2} n(\mathrm{d}\tau\mathrm{d}u)\right].$$

§18 无穷可分分布的典范形

令 $\varPhi$ 为无穷可分分布. 如在 §16 里所证明的那样, $\varPhi$ 可以看做是某

些依概率连续的可加过程 $x_t(0 \leqslant t \leqslant 1)$ 的 x_1 的分布. 由于采用可分修正, 就可以假定 x_t 为可分. 利用前节的结果便得

$$x_{t+0} = y_t + \lim_n \int_{|u|>1/n} \int_{\tau=a}^{t} \Big[uN(\mathrm{d}\tau \mathrm{d}u) - \frac{u}{1+u^2} n(\mathrm{d}\tau \mathrm{d}u)\Big].$$

由于 N 是 Poisson 可加系及 (y_t) 是与 N 独立的可分正态随机过程, 若设 $E(y_t) = m(t), V(y_t) = v(t)$, 就得

$$E(\mathrm{e}^{izx_{t+0}}) = \exp\Big\{im(t) - \frac{v(t)}{2}z^2 + \lim_{n\to\infty} \int_{|u|>1/n} \int_{\tau=a}^{t} \Big(\mathrm{e}^{izu} - 1 - \frac{izu}{1+u^2}\Big) n(\mathrm{d}\tau \mathrm{d}u)\Big\},$$

设 $n_t(\mathrm{d}u) = \displaystyle\int_{\tau=a}^{t} n(\mathrm{d}\tau \mathrm{d}u)$, 则得

$$E(\mathrm{e}^{izx_{t+0}}) = \exp\Big\{i\, m(t) - \frac{v(t)}{2}z^2 + \lim_{n\to\infty} \int_{|u|>1/n} \Big(\mathrm{e}^{izu} - 1 - \frac{izu}{1+u^2}\Big) n_t(\mathrm{d}u)\Big\}.$$

在这里, 若令 $t = 1$, 则左边为 $E(\mathrm{e}^{izx_1})$, 也就是 x_1 的分布 $\varPhi$ 的特征函数 $\varphi(z)$. 记 $m(1) = m, v(1) = v, n_1(E - \{0\}) = n(E)$, 于是就有

$$\varphi(z) = \exp\Big\{imz - \frac{v}{2}z^2 + \int_{-\infty}^{\infty} \Big(\mathrm{e}^{izu} - 1 - \frac{izu}{1+u^2}\Big) n(\mathrm{d}u)\Big\}, \tag{18.1}$$

此处 m 是实数, $v \geqslant 0, n$ 是测度, 而且

$$n(\{0\}) = 0, \quad \int_{-\infty}^{\infty} \frac{u^2}{1+u^2} n(\mathrm{d}u) < \infty.$$

这就是 P. Lévy 关于无穷可分分布的典范形. 反之, 对于满足上述条件的 $m, v, n(\mathrm{d}u)$, 若以 (18.1) 来定义 $\varphi(z)$, 则这就成为无穷可分分布的特征函数. 为了证明 (18.1) 是特征函数, 只须注意到, 若令使得 $\varphi(z) = \exp(\psi(z))$ 为特征函数的 $\psi(z)$ 全体为 $\boldsymbol{\varPsi}$, 则只要注意下列的事实即可:

(i) 从正态分布以及 Poisson 分布的特征函数的形状看来, 可知 imz, $-\frac{v}{2}z^2, \lambda(\mathrm{e}^{iz} - 1) \in \boldsymbol{\varPsi}$,

(ii) $\psi(z) \in \boldsymbol{\varPsi} \Rightarrow \psi(zu) \in \boldsymbol{\varPsi}$,

(iii) $\psi_1(z), \psi_2(z), \cdots, \psi_n(z) \in \boldsymbol{\Psi} \Rightarrow \sum\limits_i \psi_i(z) \in \boldsymbol{\Psi}$,

(iv) $\psi_n(z) \in \boldsymbol{\Psi}, \psi_n(z) \to \psi(z)$(广义一致)$\Rightarrow \psi(z) \in \boldsymbol{\Psi}$.

其次, 来证明 (18.1) 对应着一个无穷可分分布. 若在 (18.1) 中, 将 m, v 和 n 各乘以 $1/n$ 而得到的函数记为 $\varphi_n(z)$, 则 $\varphi(z) = \varphi_n(z)^n$, 又因当 $n \to \infty$ 时 $\varphi_n(z) \to 1$(广义一致), 所以, 若分别令对应于 $\varphi(z)$ 和 $\varphi_n(z)$ 的分布 (其存在性在前面已有了证明) 为 Φ 和 Φ_n, 则

$$\Phi = \Phi_n * \Phi_n * \cdots * \Phi_n (n\text{个}), \quad \Phi_n \to \text{单位分布},$$

于是 Φ 是无穷可分的.

这样一来, 就可知道 $\varphi(z)$ 为无穷可分分布的特征函数的必要而且充分的条件是: $\varphi(z)$ 能够写成 (18.1) 的形状. 并且还可以证明, 在这场合下, (18.1) 中的 m, v 和 $n(\mathrm{d}u)$ 能由分布所唯一确定, 但证明在这里从略.

正态分布就是在 (18.1) 里 $n(\mathrm{d}u) \equiv 0$ 的场合, 而 Poisson 分布就是这样的场合: $v = 0, n(\mathrm{d}u)$ 的支集仅是一点 1 而且 $n(\{1\}) = \lambda, m = \lambda/2$.

前面证明了无穷可分的分布可以表示为某一个可加过程的增量的分布. 不仅如此, 而且还可表为对时间齐次的可加过程的增量的分布. 实际上, 若把 Φ 的特征函数写成 (18.1) 的形状, 而且设 $\varphi(z) = \exp(\psi(z))$, 则 $\varphi_t(z) = \exp\{t\psi(z)\}$ 也是特征函数. 若令 $\Phi_{st} (0 \leqslant s \leqslant t \leqslant 1)$ 为对应于 $\varphi_{t-s}(z)$ 的分布, 则由 $\varphi_t(z)\varphi_s(z) = \varphi_{t+s}(z)$, 得出 $\Phi_{st} * \Phi_{tu} = \Phi_{su} (s \leqslant t \leqslant u)$. 因此, 存在可加过程 $x_t, 0 \leqslant t \leqslant 1$, 使得 $x_t - x_s (t \geqslant s)$ 的分布为 Φ_{st}. 因为 Φ_{st} 仅与 $t - s$ 有关, 所以 x_t 是对时间齐次的.

(18.1) 又可以写成下述的形状:

$$\varphi(z) = \exp\left\{inz + \int_{-\infty}^{\infty} \left(\mathrm{e}^{izu} - 1 - \frac{izu}{1+u^2}\right) \frac{1+u^2}{u^2} G(\mathrm{d}u)\right\}, \tag{18.2}$$

此处 $G(\mathrm{d}u)$ 是 $\mathbb{R}^1$ 上的有界测度. 当 $u = 0$ 时, 被积函数在 $u = 0$ 时是没有意义的, 但可以 $u \to 0$ 时的极限为其定义, 即为 $-z^2/2$. 所以 "0" 这一点的 G 测度 $G(0)$ 相当于 (18.1) 的 v. (18.2) 是 A. Khinchin 所获得的.

特别是 Φ 的方差 $V(\Phi)$ 为有限时, (18.1) 可以写为

$$\varphi(z) = \exp\left\{imz - \frac{v}{2}z^2 + \int_{-\infty}^{\infty} (\mathrm{e}^{izu} - 1 - izu) n(\mathrm{d}u)\right\},$$

$$\int u^2 n(\mathrm{d}u) < \infty. \tag{18.3}$$

此处的 m 与 (18.1) 中的 m 是不同的, 而 $v, n(\mathrm{d}u)$ 却是相同的. 这是 Kolmogoroff 所建立的形式. 又当

$$\int_{-1}^{1} |u| n(\mathrm{d}u) < \infty$$

时, 则可以写成

$$\varphi(z) = \exp\left\{imz - \frac{v}{2}z^2 + \int_{-\infty}^{\infty} (\mathrm{e}^{izu} - 1) n(\mathrm{d}u)\right\}. \tag{18.4}$$

这时 v, n 仍与 (18.1) 的那些相同, 而 m 就与 (18.1) 的有所不同.

特别地, 若所对应的可加过程仅凭跳跃而变化时, 则

$$\varphi(z) = \exp\left\{\int_{-\infty}^{\infty} (\mathrm{e}^{izu} - 1) n(\mathrm{d}u)\right\}, \tag{18.5}$$

又若仅凭正的跳跃而变化时, 则

$$\varphi(z) = \exp\left\{\int_{0}^{\infty} (\mathrm{e}^{izu} - 1) n(\mathrm{d}u)\right\}. \tag{18.6}$$

§19 Poisson 过程的各种构成方法

Poisson 过程原先是这样引进的: 它是在各瞬间的独立增量加起来的整数值随机过程 x_t, 而且还满足

$$\begin{aligned} &P(\Delta x_t = 0) = 1 - \lambda \cdot \Delta t + o(\Delta t), \quad P(\Delta x_t = 1) = \lambda \cdot \Delta t + o(\Delta t), \\ &P(\Delta x_t = k) = o(\Delta t), \quad k \geqslant 2. \end{aligned} \tag{19.1}$$

若对时间不是齐次的, 那么以 $\Delta\lambda(t)$ 来代替 $\lambda \cdot \Delta t$ 即可. 例如出租汽车的司机发生事故的次数, 作为第一次近似来说, 认为可以由 Poisson 过程来描述. 因为在交通量少的时候事故就少, 而在上下班时间事故就多, 所以对时间不是齐次的. 又因在发生事故后的一定时间内是会特加留心的, 所以对独立增量的假设也并不符合实际的. 若考虑到这些因素, 则应有第二近似, 但现在不去讨论它.

严格地说, 所谓独立增量也就是可加过程. 故在对时间齐次的场合下, 可由 (18.6) 得

$$\varphi_{ts}(z) = E(\mathrm{e}^{iz(x_t - x_s)}) = \exp\left\{(t-s)\int_0^{\infty}(\mathrm{e}^{izu}-1)n(\mathrm{d}u)\right\}.$$

又因跃度是正整数, 所以 $n(\mathrm{d}u)$ 的支集是 $\{1,2,3,\cdots\}$, 而且

$$\begin{aligned}\varphi_{st}(z) &= \exp\left\{(t-s)\sum_k(\mathrm{e}^{ikz}-1)c_k\right\}, \quad \sum_{k=1}^{\infty}c_k = \int_1^{\infty}n(\mathrm{d}u) < \infty,\\ &= \exp\left\{1-\left(\sum_k c_k\right)(t-s)+(t-s)\sum_k \mathrm{e}^{ikz}\cdot c_k + o(t-s)\right\},\end{aligned}$$

因而由 (19.1) 得 $c_2 = c_3 = \cdots = 0$, 且得

$$\varphi_{st}(z) = \exp((t-s)(\mathrm{e}^{iz}-1)c_1).$$

这就是说, $x_t - x_s$ 的分布实际就是 Poisson 分布. 对时间不齐次的场合也可以做出同样论证.

令 $x_t, t\in[0,\infty)$, 为齐次的可分 Poisson 过程. 令 $x_t(\omega)$ 的样本过程的增加时刻为 $T_0(\omega), T_0(\omega)+T_1(\omega), T_0(\omega)+T_1(\omega)+T_2(\omega),\cdots$. $T_0(\omega), T_1(\omega), T_2(\omega),\cdots$ 是在 $0,1,2,\cdots$ 上的停留时间. 每个 $T_0, T_1, T_2,\cdots$ 都具有相同分布 (叫做**指数分布**exponential distribution), 即

$$P\{\omega/T_n > t\} = \mathrm{e}^{-\lambda t}, \quad n = 0,1,\cdots,$$

而且是相互独立的. 首先, 对于 T_0 得

$$P\{\omega/T_0 > t\} = P\{\omega/x_t = 0\} = \mathrm{e}^{-\lambda t}.$$

故 T_0 的分布密度是 $\lambda\mathrm{e}^{-\lambda t}\mathrm{d}t$. 其次, 令 $t_1 > t_0$, 而求 (T_0, T_1) 的分布, 得到

$$\begin{aligned}&P\{\omega/T_0 > t_0, T_0+T_1 > t_1\}\\ &= P\{\omega/x_{t_0} = 0, x_{t_1} = 0\text{或者}1\}\\ &= P\{\omega/x_{t_0} = 0, x_{t_1} = 0\} + P\{\omega/x_{t_0} = 0, x_{t_1} = 1\}\\ &= \mathrm{e}^{-\lambda t_1} + \mathrm{e}^{-\lambda t_0}\mathrm{e}^{-\lambda(t_1-t_0)}\cdot\lambda(t_1-t_0)\\ &= \mathrm{e}^{-\lambda t_1}[1+\lambda(t_1-t_0)],\end{aligned}$$

$$\begin{aligned}E\{f(T_0,T_1)\} &= E\{f(T_0, T_0+T_1-T_0)\}\\ &= \iint_{t_1>t_0} f(t_0, t_1-t_0)\frac{\partial^2(\mathrm{e}^{-\lambda t_1}(1+\lambda(t_1-t_0)))}{\partial t_1\partial t_0}\mathrm{d}t_0\mathrm{d}t_1\end{aligned}$$

$$= \iint\limits_{t_1>t_0} f(t_0, t_1 - t_0)\lambda^2 \mathrm{e}^{-\lambda t_1} \mathrm{d}t_0 \mathrm{d}t_1$$

$$= \int_{t_1=0}^{\infty} \int_{t_0=0}^{\infty} f(t_0, t_1)\lambda^2 \mathrm{e}^{-\lambda(t_0+t_1)} \mathrm{d}t_0 \mathrm{d}t_1.$$

故 (T_0, T_1) 的分布密度是 $\lambda \mathrm{e}^{-\lambda t_0} \cdot \lambda \mathrm{e}^{-\lambda t_1}$. 由此知道, T_0 和 T_1 的分布是相同的指数分布, 而且两变量是独立的. 对于 $T_0, T_1, \cdots, T_n$ 也是一样的.

利用这种停留时间的关系, 就可以用下述方式来构造 Poisson 过程. 令 $T_0, T_1, T_2, \cdots$ 为独立随机变量序列, 且设

$$P\{\omega / T_n > t\} = \mathrm{e}^{-\lambda t}, \quad n = 0, 1, 2, \cdots.$$

若设

$$x_t = \inf\{n / T_0 + T_1 + \cdots + T_n > t\},$$

则 x_t 为 Poisson 过程. 下面举出一些实例来说明这种构造方法.

现在假设电灯泡的耐久时间服从指数分布. 也就是假设在 t 时间和 $(t + \mathrm{d}t)$ 时间之间烧掉的概率为 $\lambda \mathrm{e}^{-\lambda t} \mathrm{d}t$. 若使用这种电灯泡, 且在它烧毁后便立即换上新的, 那么, 在 t 时间之内烧掉的电灯泡的个数 x_t 就为 Poisson 过程. 这种构成法在对时间不齐次的场合是不能适用的.

另一个有趣的构成法如下. 以服从 Poisson 分布的随机变量 x 和服从 [0,1] 上的均匀分布而且相互独立的随机变量 $y_1, y_2, \cdots$ 为基础. x 与 $y_1, y_2, \cdots$ 也假定是独立的. 又令 $[0, t]$ 的示性函数为 $c_t(\xi)$. 若设

$$x_t = \sum_{i=1}^{x} c_t(y_i),$$

则 x_t 为 Poisson 过程. 实际上, 若取 $0 = t_0 < t_1 < \cdots < t_n = 1$, 又考虑 $(x_{t_i} - x_{t_{i-1}}, i = 1, 2, \cdots, n)$ 的分布时, 则得

$$\begin{aligned}
&P\{\omega / x_{t_i} - x_{t_{i-1}} = k_i, i = 1, 2, \cdots, n\} \\
&= P\left\{\omega / x = \sum k_i, \text{在} y_1, \cdots, y_n \text{之中有} k_i \text{个落入} [t_{i-1}, t_i], i = 1, 2, \cdots, n\right\} \\
&= \mathrm{e}^{-\lambda} \cdot \frac{\lambda^k}{k!} \frac{k!}{k_1! \cdots k_n!} \prod_i (t_i - t_{i-1})^{k_i}, \quad k = \sum k_i, \\
&= \prod_i \mathrm{e}^{-\lambda(t_i - t_{i-1})} \frac{(\lambda(t_i - t_{i-1}))^{k_i}}{k_i!}.
\end{aligned}$$

这就是说, $x_t, 0 \leqslant t \leqslant 1$, 是 Poisson 过程. 在对时间不齐次的场合, 也可以同样构成.

§20 复合 Poisson 过程

试考虑仅凭跳跃而变化并对时间齐次的可加过程 $x_t, 0 \leqslant t < \infty$. 显然

$$\varphi_{st}(z) \equiv E(\mathrm{e}^{iz(x_t - x_s)}) = \exp\left\{(t-s)\int_{-\infty}^{\infty}(\mathrm{e}^{izu}-1)n(\mathrm{d}u)\right\}. \tag{20.1}$$

这时 n 满足

$$\int_{|u|>1} n(\mathrm{d}u) < \infty, \quad \int_{|u|\leqslant 1} |u| n(\mathrm{d}u) < \infty. \tag{20.2}$$

此时, 高度的绝对值大于某个正数的跳跃, 在有限时间内就具概率为 1 地为有限个, 但高度接近于 0 的跳跃, 一般地说是无限的. 但因跳跃高度的和具概率为 1 地是绝对收敛的, 所以 $x_t - x_s$ 就为跳跃高度的和所确定. 若考虑比 (20.2) 稍强的条件, 而设

$$\lambda = \int_{-\infty}^{\infty} n(\mathrm{d}u) < \infty \tag{20.3}$$

时, 则在有限时间内的跳跃具概率为 1 地为有限个. 这样的可加过程叫做复合 Poisson 过程.

现在, 若设

$$\Phi(\mathrm{d}u) = \lambda^{-1} n(\mathrm{d}u),$$

则 Φ 为实数轴上的概率分布. 把 (20.1) 改写一下, 就得

$$\varphi_{st}(z) = \exp\left\{\int_s^t \int_{-\infty}^{\infty} (\mathrm{e}^{izu}-1)\Phi(\mathrm{d}u)\cdot\lambda\mathrm{d}\tau\right\}.$$

由此式可知, 在时间 $\mathrm{d}\tau$ 之间发生跳跃的概率是

$$\int_{u=-\infty}^{\infty} \Phi(\mathrm{d}u)\lambda\cdot\mathrm{d}\tau = \lambda\cdot\mathrm{d}\tau,$$

而其跳跃高度处于 $\mathrm{d}u$ 内的概率就是 $\Phi(\mathrm{d}u)\cdot\lambda\cdot\mathrm{d}\tau$. 故 $\Phi(\mathrm{d}u)$ 就可以看作跳跃高度的分布.

现在若令 x_t 在 $[0,t]$ 内发生跳跃的次数为 N_t, 那么 N_t 就是 Poisson 过程, 而且

$$\psi_{st}(z)=E(\mathrm{e}^{iz(N_t-N_s)})=\exp\{(t-s)\lambda(\mathrm{e}^{iz}-1)\}.$$

同时, x_t 和 N_t 具有相同的跳跃时刻. 所以 N_t 的跳跃高度虽然恒等于 1, 但 x_t 的跳跃高度却随 Φ 而分布.

把 $\varphi_{st}(z)$ 的表示式改写一下, 就得

$$\begin{aligned}\varphi_{st}(z)&=\mathrm{e}^{-\lambda(t-s)(1-\phi(z))}\\&=\mathrm{e}^{-\lambda(t-s)}\sum\frac{(\lambda(t-s))^n}{n!}\phi(z)^n,\quad \phi(z)=\int\mathrm{e}^{izu}\Phi(\mathrm{d}u).\end{aligned}$$

因此, x_t-x_s 的分布 Φ_{st} 就为

$$\begin{aligned}\Phi_{st}&=\sum_n\mathrm{e}^{-\lambda(t-s)}\frac{(\lambda(t-s))^n}{n!}\Phi^{*n}\qquad(20.4)\\&=\sum_n P(n;\lambda(t-s))\cdot\Phi^{*n},\quad \Phi^{*n}=\Phi*\cdots*\Phi(n\text{个}).\end{aligned}$$

从这个观点就可以构成如下的复合 Poisson 过程. 令 $T_1,T_2,\cdots,u_1,u_2,\cdots$ 为独立随机变量, 且设

$$P\{\omega/T_n>t\}=\mathrm{e}^{-\lambda t},\quad P\{\omega/u_n\in\mathrm{d}u\}=\Phi(\mathrm{d}u).$$

对此, 若定义 x_t 为

$$x_t=u_1+u_2+\cdots+u_n\quad(T_1+\cdots+T_n\leqslant t<T_1+\cdots+T_{n+1}),(n=1,2,\cdots),$$

则这就是上述的**复合 Poisson 过程**. 此处 $T_1,T_2,\cdots$ 是到下一个跳跃发生为止的等待时间, 而 $u_1,u_2,\cdots$ 是跳跃的高度.

作为复合 Poisson 过程的例子, 可以举出出租汽车的司机由于事故而造成损失的总计. 前面已经说过, 事故次数的总计是 Poisson 过程 (假设为齐次的), 但若再令事故所造成的损失的分布为 Φ, 则损失的总计的分布可由 (20.4) 给出.

§21　稳定分布和稳定过程

若两个分布 Φ 和 Ψ 可由某个正数 $\lambda>0$ 联结成下述的关系式

$$\Psi(E) = \Phi(\lambda E), \quad \lambda E = \{\lambda \cdot \xi; \xi \in E\}, \tag{21.1}$$

则 Φ 和 Ψ 叫做同构. 若令 Φ, Ψ 的分布函数和特征函数分别为 F, G 和 φ, ψ, 则 (21.1) 等价于

$$G(x) = F(\lambda x) \tag{21.1'}$$

或者

$$\psi(\lambda z) = \varphi(z). \tag{21.1''}$$

随机变量 X 的分布和 $\lambda \cdot X(\lambda > 0)$ 的分布是同构的. 均值为 0 的正态分布 $N(\cdot; 0, v)(v > 0)$ 和标准正态分布 $N(\cdot; 0, 1)$ 是同构的.

若与分布 Φ 同构的任意两个分布 Φ_1 和 Φ_2 的卷积 $\Phi_1 * \Phi_2$ 又与 Φ 同构时, 则 Φ 叫做稳定的 (stable). 若 Φ 是稳定的, 则与 Φ 同构的分布也是稳定的. 标准正态分布就是稳定的. 对稳定的 Φ 的特征函数 $\varphi(z)$, 能得到下面的特性: 对任意的 $\lambda_1, \lambda_2 > 0$, 存在 $\lambda = \lambda(\lambda_1, \lambda_2) > 0$, 使得

$$\varphi(\lambda z) = \varphi(\lambda_1 z)\varphi(\lambda_2 z). \tag{21.2}$$

由此可以证明下述定理.

定理 1 稳定分布是无穷可分的.

事实上, 由 (21.2) 便知存在 $a > 0$, 使得 $\varphi(az) = \varphi(z)^2$. 由此可得

$$\varphi(a^n z) = \varphi(z)^{2^n}. \tag{21.3}$$

若 $a = 1$, 则 $\varphi(z) = \varphi(z)^2$. 故 $\varphi(z) = 1$ 或者 0. 因为 $\varphi(z)$ 是连续的, 而且 $\varphi(0) = 1$, 所以 $\varphi(z) \equiv 1$. 故 Φ 是单位分布, 显然是无穷可分的. 若 $a < 1$, 则 $|\varphi(z)| \equiv 1$. 这是因为, 若有一 z, 使得 $|\varphi(z)| < 1$, 那么在 (21.3) 中令 $n \to \infty$, 就得 1=0, 因而得出矛盾. 由 $|\varphi(z)| \equiv 1$ 得出 $\delta(\Phi) = 0$, 也就是 Φ 成为 δ 分布. 因而也是无穷可分的. 若 $a > 1$, 则在 (21.3) 中以 z/a^n 代替 z 就得

$$\varphi(z) = \varphi(a^{-n} z)^{2^n}. \tag{21.4}$$

因为 $\varphi(a^{-n}z)$ 也是特征函数, 而且当 $n \to \infty$ 时广义一致收敛于 1, 所以由 (21.4) 就显示了 Φ 是无穷可分的.

由上面的定理, 可知 $\varphi(z)$ 可以表示为形状

$$\varphi(z) = \mathrm{e}^{\psi(z)}, \quad \psi(z) = imz - \frac{v}{2}z^2 + \int \Big(\mathrm{e}^{izu} - 1 - \frac{izu}{1+u^2}\Big) n(\mathrm{d}u), \tag{21.5}$$

而 (21.2) 就变成

$$\psi(\lambda z) = \psi(\lambda_1 z) + \psi(\lambda_2 z). \tag{21.6}$$

由此就得下述定理.

定理 2

$$\psi(z) = \Big(-c_0 + i\frac{z}{|z|}c_1\Big)|z|^{\alpha} \quad (c_0 \geqslant 0, -\infty < c_1 < \infty, \alpha > 0).$$

在 $\psi(z) \equiv 0$ 的情况, 设 $c_0 = c_1 = 0$, 定理就成立, 因此不考虑这种情况. 由 (21.6) 可知, 对正整数 n 存在 $a_n > 0$, 使得

$$\psi(a_n z) = n\psi(z). \tag{21.7}$$

至于对正的有理数 $r = q/p$ 来说, 若设 $a_r = a_n/a_m$, 则

$$\psi(a_r z) = r\psi(z). \tag{21.7'}$$

现在来证 a_r 是由 r 唯一确定的. 为此, 只要由假设 $\psi(\alpha z) = \psi(\beta z), \alpha > \beta > 0$, 而引出矛盾就行了. 设 $\gamma = \beta/\alpha$, 就得 $\psi(z) = \psi(\gamma z) = \psi(\gamma^2 z) = \cdots = \psi(\gamma^n z) \to \psi(0) = 0$. 故得 $\psi(z) \equiv 0$, 而这就是除外的场合. 由 (21.7′) 可知, 对正的有理数 r, s 来说, $\psi(a_r a_s z) = r\psi(a_s z) = rs\psi(z) = \psi(a_{rs} z)$. 故 $a_r a_s = a_{rs}$. 又由 (21.7′) 得

$$\psi(z) = r^n \psi(a_r^{-n} z).$$

当 $r \leqslant 1$ 时, 若 $a_r > 1$, 则 $\psi(z) \to \psi(0) = 0$, 于是得出矛盾. 因而若 $r \leqslant 1$, 则 $a_r \leqslant 1$. 若 $r \leqslant s$, 则 $a_r = a_{r/s} a_s \leqslant a_s$. 故只要 r 在有界区间里变动, a_r 也总是有界的. 若 $r \to 1$ 则得 $a_r \to 1$. 这是因为, 若令 a_r 的某个子序列接近于 a, 那么由上面的讨论便知 a 是有限的. 若把 (21.7′) 中的序列代以这个子序列, 则得其极限 $\psi(az) = \psi(z)$. 故 $a = 1$. 由以上所述便知, a_r 在有界区间上关于 r 是一致连续的. 故若对正的实数 t, 定义 $a_t = \lim\limits_{r \to t} a_r$($r$ 是有理数) 时, 则得

$$\psi(a_t z) = t\psi(z), \quad a_t a_u = a_{tu}, \quad a_t \text{为连续}.$$

故由最后的两个条件就得出 $a_t = t^{1/\alpha}(\alpha > 0)$ 和

$$\psi(az) = a^\alpha \psi(z).$$

因而设 $z = 1$ 就得 $\psi(a) = a^\alpha \psi(1)$. 并且 $\psi(-a) = \overline{\psi(a)} = a^\alpha \overline{\psi}(1)$. 故

$$\psi(z) = |z|^\alpha \Big(-c_0 - ic_1 \frac{z}{|z|} \Big).$$

因为 $|\varphi(z)| \leqslant 1$, 所以 c_0 必需为非负. 由于 $\varphi(z) = \exp \psi(z)$ 是无穷可分的分布的特征函数, 所以可以定义对时间齐次的可加过程 $x_t, 0 \leqslant t < \infty$, 使得

$$\varphi_{ts}(z) = E(\mathrm{e}^{iz(x_s - x_t)}) = \exp\{(s-t)\psi(z)\}.$$

若假设 $\psi(z)$ 能够定成 (21.5) 的形状, 则 x_t 从 0 到 t 之间的跳跃高度中属于 $\mathrm{d}u$ 的个数的均值为 $t \cdot n(\mathrm{d}u)$. 又对正数 a, 若令 $y_t = ax_t$, 则它也是可加过程, 而且

$$\varphi'_{ts}(z) \equiv E(\mathrm{e}^{iza(x_s - x_t)}) = \exp\{(s-t)\psi(az)\} = \exp\{(s-t)a^\alpha \psi(z)\},$$

因此, y_t 从 0 到 t 之间的跳跃高度中属于 $\mathrm{d}u$ 的个数的均值就为 $t{\cdot}a^\alpha{\cdot}n(\mathrm{d}u)$. 由于 $y_t = a \cdot x_t$, 所以后者又等于 $tn(\mathrm{d}u/a)$, 即 x_t 从 0 到 t 的跳跃高度中属于 $\mathrm{d}u/a$ 的均值. 故

$$n(\mathrm{d}u/a) = a^\alpha \cdot n(\mathrm{d}u),$$

由此得

$$n_+(x) = \int_x^\infty n(\mathrm{d}u) = \int_1^\infty n(x \cdot \mathrm{d}u) = \int_1^\infty x^{-\alpha} n(\mathrm{d}u) = x^{-\alpha} n_+(1),$$

且有 $n(\mathrm{d}u) = c \cdot u^{-\alpha-1}\mathrm{d}u (u > 0)$, 此处 c 为常数. 当 $u < 0$ 时也可得同样的结果.

这时, 可设

$$\begin{aligned} n(\mathrm{d}u) &= c_+ \cdot u^{-\alpha-1}\mathrm{d}u \quad (u > 0) \\ &= c_- |u|^{-\alpha-1}\mathrm{d}u \quad (u < 0) \quad (c_\pm \geqslant 0). \end{aligned}$$

对于 α 的限制, 则有下述的定理.

定理 3 Φ 是正态分布, 此外必需 $0 < \alpha < 2$.

证明 若 $c_+ = c_- = 0$, 则 Φ 是正态分布.若 c_+ 和 c_- 之中有一个是正的, 则由于 $\int_{-1}^{1} u^2 n(\mathrm{d}u) < \infty$ 而必需 $0 < \alpha < 2$.

将 α 的值分为 $0 < \alpha < 1, 1 < \alpha < 2$ 和 $\alpha = 1$ 三种情形来讨论.

(i) $0 < \alpha < 1$. 这时因为

$$\int_{-\infty}^{\infty} n(\mathrm{d}u) = \infty, \quad \int_{-\infty}^{\infty} |u| n(\mathrm{d}u) < \infty,$$

所以得出

$$\psi(z) = imz - \frac{v}{2} z^2 + c_+ \int_0^{\infty} (\mathrm{e}^{izu} - 1) \frac{\mathrm{d}u}{u^{\alpha+1}} + c_- \int_{-\infty}^{0} (\mathrm{e}^{izu} - 1) \frac{\mathrm{d}u}{|u|^{\alpha+1}}.$$

上式中含有积分的项应为 $O(|z|^\alpha)$(当 $z \to 0$ 及 $z \to \infty$ 时). 例如当 $z > 0$ 时

$$\int_0^{\infty} (\mathrm{e}^{izu} - 1) \frac{\mathrm{d}u}{u^{\alpha+1}} = z^\alpha \mathrm{e}^{-\pi i \alpha/2} \int_0^{\infty} (\mathrm{e}^{-v} - 1) \frac{\mathrm{d}v}{v^{\alpha+1}} \quad (z > 0).$$

由于 $\psi(z) = O(|z|^\alpha)$, 所以考虑 $z \to \infty$ 时的阶数, 必需 $m = v = 0$. 故

$$\psi(z) = c_+ \int_0^{\infty} (\mathrm{e}^{izu} - 1) \frac{\mathrm{d}u}{u^{\alpha+1}} + c_- \int_{-\infty}^{0} (\mathrm{e}^{izu} - 1) \frac{\mathrm{d}u}{|u|^{\alpha+1}}.$$

因此, 在这种情况下, 对应的可加过程 x_t 的样本过程具概率为 1 地是仅凭跳跃而变化的纯粹不连续函数. 并且除 $c_+ = c_- = 0$ 的特别情况以外, 跳跃的个数具概率为 1 地是无限的. 在定理 2 中已经证明 $c_0 \geqslant 0$, 但若 $c_0 = 0$, 则 $|\varphi(z)| = 1$, 因而 Φ 就成为 δ 分布, 并且 $\psi(z) = imz$(m 是常数), 但这只有在 $c_1 = n = 0$ 的场合即 Φ 是单位分布的场合才能成立的 (应注意到 $0 < \alpha < 1$). 又因 $|\varphi(z)| = \exp(-c_0 |z|^\alpha)(0 < \alpha < 1)$, 所以 $\varphi(z)$ 即属于 $L_1(R^1)$, 又属于 $L_2(R^1)$. 故其 Fourier 变式是连续函数. 这就表示了与 φ 对应的分布 Φ 是具有连续密度的. 既然样本过程是纯粹不连续的, 可是分布却又具有连续密度, 这倒是有趣的现象. 这是因为不连续的函数被平均后而变为连续的缘故.

(ii) $1 < \alpha < 2$. 这时

$$\int_{-\infty}^{\infty} n(\mathrm{d}u) = \infty, \quad \int_{-\infty}^{\infty} u n(\mathrm{d}u) = \infty, \quad \int_{|u|>1} |u| n(\mathrm{d}u) < \infty.$$

因此

$$\psi(z)=imz-\frac{v}{2}z^2+c_+\int_0^{\infty}(\mathrm{e}^{izu}-1-izu)\frac{\mathrm{d}u}{u^{\alpha+1}}+c_-\int_{-\infty}^{0}(\mathrm{e}^{izu}-1-izu)\frac{\mathrm{d}u}{|u|^{\alpha+1}}.$$

因为积分项等于 $O(|z|^{\alpha})$, 所以考虑到 $z\to 0$ 及 $z\to\infty$ 时的阶数, 就得 $m=v=0$, 而且

$$\psi(z)=c_+\int_0^{\infty}(\mathrm{e}^{izu}-1-izu)\frac{\mathrm{d}u}{u^{\alpha+1}}+c_-\int_{-\infty}^{0}(\mathrm{e}^{izu}-1-izu)\frac{\mathrm{d}u}{|u|^{\alpha+1}}.$$

对应的可加过程的跳跃个数以及高度的绝对值之和都具概率为 1 地为 ∞, 但是任意取 $\beta(>\alpha)$ 时, 跳跃高度的绝对值的 β 乘幂之和

$$\sum_{0<\tau<t}|x_{\tau+0}-x_{\tau-0}|^{\beta}$$

却以概率为 1 地为有限. 由于其中高度的绝对值大于 1 的跳跃是有限的, 因此, 只要考虑在 1 以下的绝对值就行了. 因为其均值为

$$\int_0^t\int_{-1}^{1}|u|^{\beta}\frac{\mathrm{d}u}{u^{\alpha+1}}\mathrm{d}\tau<\infty.$$

所以其本身就是具概率为 1 地为有限.

(iii) $\alpha=1$. 因为

$$\psi(z)=ic_1z-c_0|z|$$

又

$$\int_{-\infty}^{\infty}\left(\mathrm{e}^{izu}-1-i\frac{zu}{1+u^2}\right)\frac{\mathrm{d}u}{u^2}=2\int_0^{\infty}(\cos zu-1)\frac{\mathrm{d}u}{u^2}=-\pi|z|,$$

所以

$$\psi(z)=ic_1z-\frac{c_0}{\pi}\int_{-\infty}^{\infty}\left(\mathrm{e}^{izu}-1-i\frac{zu}{1+u^2}\right)\frac{\mathrm{d}u}{u^2}.$$

这时对应的分布是 Cauchy 分布. 因而对应的可加过程叫做 **Cauchy 过程**(Cauchy Process).

如上面得出的可加过程那样, 对应于稳定分布的就叫做**稳定过程**(stable process). (i), (ii) 的场合和 (iii) 的 Cauchy 过程, 正态随机过程 (见定理 3), 以及作为后者的退化的情形而有 $x_t=m\cdot t$(m 是常数) 的场合, 等等, 都是稳定过程.

第3章　平稳过程

§22　平稳过程的定义

所谓平稳过程是指描述对时间的变化保持着平衡性的现象的随机过程. 平衡的定义有强弱两种. 令 $x_t(\omega), t \in T = (-\infty, \infty)$, 为一随机过程, 且设

$$m(t) = E(x_t), \quad v(t,s) = E((x_t - E(x_t))(x_s - E(x_s))),$$
$$\varPhi_{t_1,t_2,\cdots,t_n}(E) = P\{\omega/(x_{t_1}, x_{t_2}, \cdots, x_{t_n}) \in E\}.$$

若对任意的 t, s 和 h 均有

$$m(t+h) = m(t), \quad v(t+h, s+h) = v(t,s),$$

则称 x_t 为**弱平稳过程**(weakly stationary stochastic process). 此时, $m(t)$ 是常数 (m), 而 $v(t,s)$ 是 $t-s$ 的函数 $(v(t-s))$. 若对任意的 n 和 $\{t_i\}$, 均有

$$\varPhi_{t_1+h,t_2+h,\cdots,t_n+h} = \varPhi_{t_1,t_2,\cdots,t_n},$$

则称 x_t 为**强平稳过程**(strongly stationary stochastic process).

定理 1　若 x_t 是强平稳过程, 并且 $E(|x_0|^2) < \infty$, 则 $x_t, t \in T$, 亦为弱平稳过程.

证明　由强平稳性可知 $E(|x_t|^2) = E(|x_0|^2) < \infty$, 故 $m(t), v(t,s)$ 皆存在. 又因

$$m(t) = \int \xi \varPhi_t(\mathrm{d}\xi),$$
$$v(t,s) = \iint (\xi - m(t))(\eta - m(s)) \varPhi_{ts}(\mathrm{d}(\xi,\eta)),$$

由此及过程的强平稳性即可推出弱平稳性.

其逆虽然不一定成立, 但对于正态随机过程来说, 成立着下述的定理.

定理 2　若正态随机过程 x_t 是弱平稳的, 那么它亦为强平稳.

证明　若任意取 $t_1, t_2 \cdots, t_n$, 并设

$$M = (m(t_i)), \quad V = (v(t_i, t_j)),$$

则由正态性的假定可知 $\Phi_{t_1 \cdots t_n}$ 就是 $N(\cdot; M, V)$. 由弱平稳性可知, 以 t_i+h 代替 t_i 后 M 和 V 仍不改变, 因而 $\Phi_{t_1+h, t_2+h, \cdots, t_n+h} = \Phi_{t_1 t_2 \cdots t_n}$. 这就是说, x_t 是强平稳的.

若以整数集合 $\{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}$ 来代替 $(-\infty, \infty)$, 则上述的各种结论仍然成立. 这时, x_t 就叫做**平稳序列** (stationary random sequence).

在 $x_t(\omega)$ 取复数值的场合 (复随机过程) 下, 也可以定义平稳性. 它与实值场合的不同之处就在于以

$$v(t, s) = E((x_t - m(t))\overline{(x_s - m(s))}) \quad (\bar{\xi} = \xi\text{的共轭复数})$$

来代替实值场合的 $v(t, s)$. 定理 1 当然仍旧成立. 相当于定理 2 的定理也成立, 不过为此需要引进复正态分布这一概念 (见 §27 和 §28). 以后, 若无特别声明时, 所谓平稳过程皆指复平稳过程.

§23 关于研究平稳过程的准备知识

下面简述平稳过程论所用到的知识.

Bochner 定理 这在叙述特征函数的时候已经有了说明, 此处以略有不同的形式再叙述一遍. 若 $\varphi(t)$ 为一定义在 $(-\infty, \infty)$ 上的复数值函数且满足下述条件:

(i) 正定: $\sum\limits_{ij} \varphi(t_i - t_j)\xi_i\bar{\xi}_j \geqslant 0$,

(ii) 连续: 当 $t \to 0$ 时, $\varphi(t) \to \varphi(0)$,

则可以确定一个有界、右连续、非降的函数 $F(\lambda)$, 使得

$$\varphi(t) = \int_{\mathbb{R}^1} \mathrm{e}^{-i2\pi t\lambda} \mathrm{d}F(\lambda), \quad F(-\infty) = 0.$$

此处, 在指数里加上 2π 仅仅是为了方便而已.

Stone 定理 [①] 设 $U_t, -\infty < t < \infty$, 为一族 Hilbert 空间 $\boldsymbol{H}$ 的酉算子, 且满足

① 参见关肇直编的《泛函分析讲义》, 高等教育出版社, 1958, 第 3 章, §2 ~ §4.
—— 校者注

(i) 连续: $(U_t f, g)$ 关于 t 连续 (或者可测也可以),

(ii) 群: $U_t U_s = U_{t+s}$,

则 U_t 存在着**谱分解**

$$U_t = \int \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}E(\lambda).$$

又若 U 是酉算子, 则存在谱分解

$$U = \int \mathrm{e}^{-i2\pi\lambda}\mathrm{d}E(\lambda).$$

遍历定理(ergodic theorem) 令 $\Omega(\boldsymbol{B}, P)$ 为一概率空间, S 为从 Ω 映到 Ω 自身的一对一的保测变换. 此处所谓**保测**(measure preserving) 是指: 若 E 是可测的, 则 $SE, S^{-1}E$ 也都可测, 而且 $P(SE) = P(S^{-1}E) = P(E)$. 这时, 就 $f \in L^1(\Omega)$ 来说, 对于几乎所有的 ω, 有下列等式成立:

$$f^*(\omega) = \lim_{\substack{n\to\infty \\ m\to-\infty}} \frac{1}{n-m}[f(S^{m+1}\omega) + f(S^{m+2}\omega) + \cdots + f(S^n\omega)].$$

并且 $f^*(S\omega) = f^*(\omega)$(a.e.). 这叫做 G. D. Birkhoff 的**个体遍历定理**(individual ergodic theorem).

假设**广义函数**(distribution) 的理论 ① 已为读者所知. 除了取复数值的普通的广义函数以外, 还考虑取值于 Hilbert 空间里的广义函数. 其定义与普通的场合是完全一样的.

若令 $\Omega(\boldsymbol{B}, P)$ 为概率空间, 则依照普通的方法, $\boldsymbol{H} = L^2(\Omega)$ 就可以看做 Hilbert 空间. 若用 Hilbert 空间的语言来表达概率论在 $\Omega(\boldsymbol{B}, P)$ 上的概念, 于是就可记为

$$E(x) = (x, 1),$$
$$E(x\cdot\overline{y}) = (x, y), \quad \text{特别地}, E(|x|^2) = \| x \|^2,$$
$$x_n \to x(\text{平均收敛}) \Leftrightarrow \| x_n - x \| \to 0.$$

若 $x_t(\in \boldsymbol{H})$ 关于 t 连续 (在范数的意义下), 又依范数为有界, 而且当 $f(t) \in L^1(\mathbb{R}^1)$ 时, 则在范数收敛的意义下, 可以定义 $\int f(t)x_t\mathrm{d}t$. 这对 t 的变动范围是区间的场合也适用. 其次, 定义

$$I(f) = \int f(t)\mathrm{d}y_t.$$

① 例如可参见冯康编的《广义函数论》, 数学进展, 3(1), 1955, 405~590. —— 校者注

特别重要的就是当 y_t 具有**正交增量性**(orthogonal increment) 的场合. 即对任意互不相交的区间 $(t_1, t_2]$ 和 $(s_1, s_2]$, $y_{t_2} - y_{t_1}$ 和 $y_{s_2} - y_{s_1}$ 是正交的. 这时, 就存在单调增函数 $F(t)$(除去可加常数外为唯一确定), 使得

$$F(s) - F(t) = \| y_s - y_t \|^2 \quad (s > t).$$

特别若假定 $F(t)$ 为右连续, 则 y_t 也为右连续. 由 F 而确定的 Lebesgue-Stieltjes 测度仍以同一符号 F 来表示, 即

$$F(E) = \int_E \mathrm{d}F(t).$$

现往证: 对 $f \in L^2(\mathbb{R}^1, F)$, 可确定上述的 $I(f)$. 首先, 若 f 为有界阶梯函数 (记其全体为 $\boldsymbol{J}$), 则 $I(f)$ 的定义与平常的一样. 又因

$$\| I(f) \| = \| f \| \quad (\| f \| \text{是} L^2(\mathbb{R}^1, F) \text{上的范数}),$$

故可利用普通的方法, 将定义在 $\boldsymbol{J}$ 上的算子 $I(f)$ 扩张到 $\bar{\boldsymbol{J}}$ 即 $L^2(\mathbb{R}^1, F)$ 上去, 这就是所要证明的. 易知

$$I(\alpha f + \beta g) = \alpha I(f) + \beta I(g), \quad (If, Ig) = (f, g).$$

§24 弱平稳过程的谱分解

令 $x_t, -\infty < t < \infty$, 为一弱平稳过程, 且设

$$E(x_t) = m, \quad E((x_t - m)\overline{(x_s - m)}) = v(t - s). \tag{24.1}$$

除了 $x_t(\omega) \equiv m$ 这一特殊场合外, $v(0) > 0$, 所以, 可用 $(x_t - m)/\sqrt{v(0)}$ 来代替 x_t, 因而不妨假定

$$E(x_t) = 0, \quad E(x_t \overline{x}_s) = v(t - s), \quad E(|x_t|^2) = v(0) = 1. \tag{24.1'}$$

若将 $x_t, -\infty < t < \infty$, 视为 Hilbert 空间 $\boldsymbol{H} = L^2(\Omega)$ 中的曲线, 则由 (24.1′) 可得

$$(x_t, 1) = 0, \quad (x_t, x_s) = v(t - s), \quad \| x_t \| = 1. \tag{24.1''}$$

所以, 这曲线在 $\boldsymbol{H}$ 中既属于 $\{1\}$的正交补空间 $\boldsymbol{H}' = \{y/y \perp 1\}$, 又位于 $\boldsymbol{H}$ 中的单位球上. 而 $v(t - s)$ 就是 $\boldsymbol{H}$ 中两个向量 x_t 和 x_s 的夹角的余弦. 称它为 x_t 和 x_s 的相关系数 (correlation coefficient).

更假定 x_t 是**依范数连续的**, 即

$$\lim_{t\to s}\| x_t - x_s \|= 0. \tag{24.2}$$

这样, 由前章的 Poisson 过程的例子就可知道, 此时并不能推出样本过程的连续性, 然因

$$P\{\omega/|x_t - x_s| > \varepsilon\} \leqslant \| x_t - x_s \|^2 /\varepsilon^2,$$

所以依概率连续性可以由依范数连续性推导出来.

定理 1(A. Khinchin) $v(t)$ 具有谱分解

$$v(t) = \int \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}F(\lambda), \quad F(-\infty) = 0. \tag{24.3}$$

此处 $F(\lambda)$ 是由 $v(t)$ 确定的有界、右连续的增函数. $F(\lambda)$ 叫做 $v(t)$ 的谱函数.

证明 由 $v(t)$ 的定义可知

$$\sum_{ij} v(t_i - t_j)\xi_i\bar{\xi}_j = \left\| \sum \xi_i x(t_i)\right\|^2 \geqslant 0.$$

又由 (24.2), 得

$$v(t) = (x_t, x_0) \to (x_0, x_0) = v(0) \quad (t \to 0).$$

故由 Bochner 定理就可以定出 F.

定理 2(A. Kolmogoroff) x_t 具有谱分解

$$x_t = \int \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}y_\lambda, \quad y_{-\infty} = 0. \tag{24.4}$$

此处的 y_λ 为具有正交增量的过程, 而且

$$\| y_\lambda \|^2= F(\lambda), \tag{24.5}$$

y_λ 由 x_t 而定.

证明 令 $\{x_t\}$ 的线性组合所构成的线性流形为 $\boldsymbol{A}$, 并令 $\bar{\boldsymbol{A}}$ 为 $\boldsymbol{H}_0$.$\boldsymbol{H}_0$ 就是 $\boldsymbol{H}_1$ 的子空间. 定义 $\boldsymbol{A}$ 中的变换群 $U_t, -\infty < t < \infty$, 如下:

$$U_t\left(\sum a_i x_{t_i}\right) = \sum a_i x_{t_i+t}.$$

为了证明这个定义是有意义的, 只要证明

$$\sum a_i x_{t_i} = 0 \Rightarrow \sum a_i x_{t_i+t} = 0$$

就行了, 但这由下式即可推出:

$$\left\| \sum a_i x_{t_i} \right\|^2 = \sum a_i \overline{a}_j v(t_i - t_j) = \left\| \sum a_i x_{t_i+t} \right\|^2. \tag{24.6}$$

而上式又表明 U_t 是从 $\boldsymbol{A}$ 到 $\boldsymbol{A}$ 上的**等距变换**(isometric transformation). 因此, U_t 可以扩张为从 $\boldsymbol{H}_0$ 到 $\boldsymbol{H}_0$ 的等距变换. 又因在 $\boldsymbol{A}$ 中成立着关系式 $U_t U_s = U_{t+s}$, 所以在 $\boldsymbol{H}_0$ 上也有这个性质. 由于当 $f \in \boldsymbol{A}$ 时 $U_t f$ 关于 t 依范数连续, 所以利用 $U_t f$ 的等距性便知, 对于 $f \in \boldsymbol{H}_0$ 来说, 连续性也成立. 故由 Stone 定理就得出 U_t 的谱分解:

$$x_t = U_t x_0 = \int \mathrm{e}^{-i2\pi\lambda t} \mathrm{d}(E(\lambda)x_0) = \int \mathrm{e}^{-i2\pi\lambda t} \mathrm{d}y_\lambda.$$

由谱分解的性质便知, y_λ 具有正交增量性, 而且

$$G(\lambda) = \| y_\lambda \|^2$$

给出增的有界右连续函数 $G(\lambda)$. 又由谱分解的性质而得

$$v(t) = (x_t, x_0) = \int \mathrm{e}^{-i2\pi\lambda t} \mathrm{d}G(\lambda), \quad G(-\infty) = 0.$$

以此与 (24.3) 比较就得 $F = G$, 所以 (24.5) 成立.

其次, 若有两个 y_λ 满足 (24.4), 记其为 y_λ 和 y'_λ, 现往证 $y_\lambda = y'_\lambda$. 令 $f \in L^1(R^1)$ 的 Fourier 变换为 $\widehat{f}$, 则

$$\int \widehat{f}(\lambda)\mathrm{d}y_\lambda = \int \widehat{f}(\lambda)\mathrm{d}y'_\lambda = \int f(t)x_t \cdot \mathrm{d}t,$$

若 $g(\lambda)$ 为连续且在一有限的区间外为零, 则 $g(\lambda)$ 可以由上面的 $\widehat{f}(\lambda)$ 来一致逼近, 因而

$$\int g(\lambda)\mathrm{d}y_\lambda = \int g(\lambda)\mathrm{d}y'_\lambda.$$

若令 $(-\infty, \mu]$ 的示性函数为 $c(\lambda)$, 对 $c(\lambda)$ 定出上述的 $g(\lambda)$, 使得

$$\int |c(\lambda) - g(\lambda)|^2 \mathrm{d}F(\lambda) < \varepsilon^2,$$

于是

$$\left\| \int c(\lambda)\mathrm{d}y_\lambda - \int g(\lambda)\mathrm{d}y_\lambda \right\|^2 = \int |c(\lambda) - g(\lambda)|^2 \mathrm{d}F(\lambda) < \varepsilon^2.$$

同理可得
$$\left\|\int c(\lambda)\mathrm{d}y'_\lambda - \int g(\lambda)\mathrm{d}y'_\lambda\right\|^2 < \varepsilon^2.$$
由此便有
$$\left\|\int c(\lambda)\mathrm{d}y_\lambda - \int c(\lambda)\mathrm{d}y'_\lambda\right\| < \varepsilon.$$
令 $\varepsilon \to 0$ 就得
$$\int c(\lambda)\mathrm{d}y_\lambda = \int c(\lambda)\mathrm{d}y'_\lambda \quad 即 y_\mu = y'_\mu.$$
这样一来定理 2 就被证明了.

例 现在来考虑 $F(\lambda)$ 是纯粹不连续的这一特殊情形, 将其写成
$$F(\lambda) = \sum_{\lambda_n \leqslant \lambda} a_n.$$
此处 $\{\lambda_n\}$ 为给定的实数列, $\{a_n\}$ 是正数列, 并设 $\sum a_n < \infty$. 此时
$$v(t) = \int \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}F(\lambda) = \sum a_n \mathrm{e}^{-i2\pi\lambda_n t}.$$
就成为概周期函数, 而 x_t 的谱分解可表为下列形式:
$$x_t = \sum y_n \mathrm{e}^{-i2\pi\lambda_n t} \quad (此处\{y_n\}是正交列, 而且 \| y_n \|^2 = a_n).$$
事实上, 这只需令 $y_n = y_{\lambda_{n+0}} - y_{\lambda_{n-0}}$ 就行了.

§25 弱平稳过程的样本过程的谱分解

令 $x_t, -\infty < t < \infty$, 为弱平稳过程, 又假定满足 (24.1′), (24.2), 由 Kolmogoroff 的谱分解定理, 就得
$$x_t = \int \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}y_\lambda, \quad y_{-\infty} = 0. \tag{25.1}$$
现在试就样本过程来考察一下这个关系. 因为右边的积分是由 $L^2(\Omega)$ 中的范数收敛来定义的, 所以, 这一关系并不对所有的 ω 皆成立. 首先, 在条件
$$\| x_t - x_s \| \to 0 (t \to s), \quad \| y_\lambda - y_\mu \| \to 0 (\lambda \downarrow \mu), \tag{25.2}$$
之下, 可得到下述的定理, 其证明从略.[①]

① 参见 J. L. Doob, Stochastic Processes. 1953, 61 页. —— 校者注

定理 1 对于 x_t 和 y_λ, 存在可测的 (对二元变量 (t,ω) 或 (λ,ω) 为可测)x_t^* 和 y_λ^*, 使 $P\{x_t=x_t^*\}=1, t\in T; P\{y_\lambda=y_\lambda^*\}=1, \lambda\in\Lambda$.

又因在 x_t^* 和 y_λ^* 之间, (25.1) 显然是成立的, 故以后不妨就假定 x_t 和 y_λ 是可测的. 由关于二元变量 (t,ω)(或者 (λ,ω)) 的可测性可知, 对于几乎所有的 ω, x_t(或者 y_λ) 关于 t(或者 λ) 是可测的.

$x_t(\omega)$ 对于几乎所有的 ω 来说, 是 t 的缓增函数 (见广义函数论). 这是因为

$$E\left\{\left(\int\frac{|x_t|}{1+t^2}\mathrm{d}t\right)^2\right\}\leqslant E\left\{\int\frac{|x_t|^2}{1+t^2}\mathrm{d}t\int\frac{\mathrm{d}t}{1+t^2}\right\}=\left(\int\frac{\mathrm{d}t}{1+t^2}\right)^2<\infty.$$

故对于几乎所有的 ω,

$$\int\frac{|x_t(\omega)|}{1+t^2}\mathrm{d}t<\infty,$$

这就是说, $x_t(\omega)$ 是 t 的缓增广义函数 (其实就是缓增函数). 同理

$$E\left\{\left(\int\frac{|y_\lambda|}{1+\lambda^2}\mathrm{d}\lambda\right)^2\right\}\leqslant E\left\{\int\frac{|y_\lambda|^2}{1+\lambda^2}\mathrm{d}\lambda\int\frac{\mathrm{d}\lambda}{1+\lambda^2}\right\}$$
$$\leqslant\left(\int\frac{\mathrm{d}\lambda}{1+\lambda^2}\right)^2<\infty.$$

因而, 对于几乎所有的 ω 来说, $y_\lambda(\omega)$ 也是 λ 的缓增 (广义) 函数. 从而作为广义函数的 y_λ 的微分 $\mathrm{D}y_\lambda$ 也是缓增广义函数.

定理 2 对于几乎所有的 ω 来说, $x_t(\omega), -\infty<t<\infty$, 是广义函数 $\mathrm{D}y_\lambda(\omega)$ 的 Fourier 变换, 即

$$x(\phi)=\mathrm{D}y_\lambda(\mathfrak{F}\phi). \tag{25.3}$$

此处
$$\mathfrak{F}\phi(\lambda)=\int\mathrm{e}^{-i2\pi\lambda t}\phi(t)\mathrm{d}t.$$

证明 因为
$$E\left[\int\frac{|x_t(\omega)|^2}{1+t^2}\mathrm{d}t\right]=\int\frac{1}{1+t^2}\mathrm{d}t<\infty,$$

所以对于几乎所有的 $\omega(\omega\in\Omega_1, P(\Omega_1)=1)$,

$$\int\frac{|x_t(\omega)|^2}{1+t^2}\mathrm{d}t<\infty.$$

故当 $\omega\in\Omega_1$ 时, 若令

$$z_\lambda(\omega)=\int_{-1}^{1}x_t\frac{\mathrm{e}^{i2\pi\lambda t}-1}{i2\pi t}\mathrm{d}t+\underset{a\to\infty}{\mathrm{l.i.m}}^{①}\Big(\int_1^a+\int_{-a}^{-1}\Big)x_t\frac{\mathrm{e}^{i2\pi\lambda t}}{i2\pi t}\mathrm{d}t, \tag{25.4}$$

则得

$$x(\phi)=\mathfrak{F}\mathrm{D}z_\lambda(\phi),\quad \phi\in\mathfrak{D}. \qquad (25.5)^{②}$$

这就是说, 当 $\omega\in\Omega_1$ 时, x 和 $\mathfrak{F}\mathrm{D}z_\lambda$ 作为 $\mathfrak{D}'$ 中的元素是一致的, 又因 $P(\Omega_1)=1$, 所以作为 $\mathfrak{D}'_{\boldsymbol{H}}(\boldsymbol{H}=L^2(\Omega))$ 的元素, x 和 $\mathfrak{F}\mathrm{D}z_\lambda$ 也是一致的. 故在 $\boldsymbol{H}$ 中有

$$x(\phi)=\int\mathfrak{F}\phi(\lambda)\mathrm{d}y_\lambda=-\int(\mathfrak{F}\phi(\lambda))'y_\lambda\mathrm{d}\lambda,$$

$$x(\phi)=\mathfrak{F}\mathrm{D}z_\lambda(\phi)=-\int(\mathfrak{F}\phi(\lambda))'z_\lambda\mathrm{d}\lambda,$$

若设 $\phi=\mathfrak{F}^{-1}\psi$, 则对任意的急减函数 ψ,

$$\int\psi'y_\lambda\mathrm{d}\lambda=\int\psi'z_\lambda\mathrm{d}\lambda,\quad 故\quad \mathrm{D}y_\lambda=\mathrm{D}z_\lambda,$$

也就是对于几乎所有的 λ, 得出 $y_\lambda=z_\lambda+c, c\in\boldsymbol{H}$. 因为在 (25.5) 中可令 $z_\lambda-c$ 来代替 z_λ, 所以对于几乎所有的 λ 来说, 等式 $y_\lambda=z_\lambda$ 在 $\boldsymbol{H}$ 中成立. 再由 (25.4) 可知, 对于几乎所有的 λ, 可以取 a_n, 使得

$$z_\lambda(\omega)=\int_{-1}^{1}\ ''+\lim_{a_n\to\infty}\Big(\int_1^{a_n}+\int_{-a_n}^{-1}\Big)\ '',$$

又由 $x_t(\omega)$ 对 (t,ω) 的可测性, 所以可使 $z_\lambda(\omega)$ 对 (λ,ω) 可测. $y_\lambda(\omega)$ 对 (λ,ω) 也是可测的. 又因对于几乎所有的 λ 和对于几乎所有的 ω 来说, $y_\lambda(\omega)=z_\lambda(\omega)$, 所以由 Fubini 定理便知, 对于几乎所有的 ω 和对于几乎所有的 λ 来说, $y_\lambda(\omega)=z_\lambda(\omega)$. 故 (25.4) 可以写成 $x(\phi)=\mathfrak{F}\mathrm{D}y_\lambda(\phi)$.

---

① l.i.m. 表示均方收敛, 即 $\lim_{a\to\infty}f_a(t)=f(t)\Leftrightarrow\lim_{a\to\infty}\int|f_a(t)-f(t)|^2\mathrm{d}t=0$.

② 公式 (25.5) 表示对任意的具有限支集的无穷可微函数 $\phi(\lambda)$,

$$\int_{-\infty}^{\infty}x_t\Big(\int_{-\infty}^{\infty}\mathrm{e}^{i2\pi\lambda t}\phi(\lambda)\mathrm{d}\lambda\Big)\mathrm{d}t=-\int_{-\infty}^{\infty}Z_\lambda\phi'(\lambda)\mathrm{d}\lambda,$$

而右边的积分等于

$$-\underset{a\to\infty}{\mathrm{l.i.m}}\int_{-\infty}^{\infty}\phi'(\lambda)\Big\{\int_{-1}^{1}x_t\frac{\mathrm{e}^{i2\pi\lambda t}-1}{i2\pi t}\mathrm{d}t+\Big(\int_1^a+\int_{-a}^{-1}\Big)x_t\frac{\mathrm{e}^{i2\pi\lambda t}}{i2\pi t}\mathrm{d}t\Big\}\mathrm{d}\lambda.$$

—— 俄译本注

若按照 N. Wiener 的样式写, 则 (25.3) 变成

$$x_t = \lim_{\varepsilon\to 0}\underset{a\to\infty}{\mathrm{l.i.m}}\int_{-a}^{a} \mathrm{e}^{-i2\pi\lambda t}\frac{y_{\lambda+s}-y_{\lambda-s}}{2\varepsilon}\mathrm{d}\lambda \tag{25.3$'$}$$

§26 关于强平稳过程的遍历定理

令 $x_t(\omega), -\infty < t < \infty$, 为可测强平稳过程. 假定它的均值存在, 于是就有

$$E(|x_t|) = E(|x_0|),$$

因为 $x_t(\omega)$ 关于二元变量 (t,ω) 为可测, 所以对于几乎所有的 ω 来说, $x_t(\omega)$ 作为 t 的函数是可测的. 又因对于 $-\infty < a < b < \infty$,

$$E\left\{\int_a^b |x_t|\mathrm{d}t\right\} = \int_a^b E(|x_t|)\mathrm{d}t = E(|x_0|)(b-a) < \infty,$$

所以对于任意给定的有限区间 (a,b) 和对于几乎所有的 ω 来说,

$$\int_a^b |x_t(\omega)|\mathrm{d}t < \infty.$$

由于这个积分当 a, b 为整数的场合 (这种场合的个数是可数的) 为有限, 由此不难推出对于几乎所有的 ω, 上述积分对一切的有限区间, 均为有限.

定理 设 $E(|x_0|) < \infty$, 则对于几乎所有的 ω, 下列极限存在:

$$x^*(\omega) = \lim_{A\to\infty}\frac{1}{2A}\int_{-A}^{A} x_t(\omega)\mathrm{d}t.$$

称 $x^*(\omega)$ 为**样本平均**(sample mean).

证明 先证下述的引理.

引理 若 $\{x_n\}$ 是平稳序列, 则当 $E(|x_0|) < \infty$ 时, 下列极限对于几乎所有的 ω 皆存在:

$$x^* = \lim_{\substack{n\to\infty\\ m\to-\infty}}\frac{1}{n-m}\sum_{k=m+1}^{n} x_k.$$

证明 设 $Z = \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}$, 由 $\mathbb{R}^Z(\boldsymbol{B}^Z)$ 添加上随机向量 $\boldsymbol{x} = \prod_k x_k$ 的分布 $\varPhi$, 便得到一个概率空间 $\mathbb{R}^Z(\boldsymbol{B}^Z, \varPhi)$. 若考虑从 $\mathbb{R}^Z$ 到它本身的一对一的映照

$$T:\quad \prod_k \xi_k \to \prod_k \xi_{k+1},$$

则这是一个保测变换. 若设 f 为由 $\varXi = \prod\limits_k \xi_k \to \xi_0$ 所决定的映照, 则 $f \in L^1(\mathbb{R}^Z)$. 故对于几乎所有 (关于 $\varPhi$) 的 $\varXi$, 下列极限存在:

$$f^*(\varXi) = \lim_{\substack{n\to\infty \\ m\to-\infty}} \frac{1}{n-m} \sum_{k=m+1}^{n} f(T^k \varXi).$$

从而对于几乎所有的 ω,

$$f^*(\boldsymbol{x}) = \lim_{\substack{n\to\infty \\ m\to-\infty}} \frac{1}{n-m} \sum_{k=m+1}^{n} x_k.$$

于是引理得证.

现在转到定理的证明上去. 若设

$$y_n = \int_n^{n+1} x_t \mathrm{d}t, \quad n = \cdots, -3, -2, -1, 0, 1, 2, 3, \cdots,$$

则这是一平稳序列 (由于证明过烦, 此处从略). 由上述的引理可知

$$x^* = \lim_{n\to\infty} \frac{1}{2n} \int_{-n}^{n} x_t \mathrm{d}t = \lim_{n\to\infty} \frac{1}{2n} \sum_{k=-n}^{n-1} y_k$$

对于几乎所有的 ω 皆存在. 其次, 若 $n < A < n+1$, 则由

$$\frac{1}{2A}\int_{-A}^{A} x_t \mathrm{d}t = \frac{n}{A}\frac{1}{2n}\int_{-n}^{n} x_t \mathrm{d}t + \frac{1}{2A}\int_{-A}^{-n} x_t \mathrm{d}t + \frac{1}{2A}\int_{n}^{A} x_t \mathrm{d}t$$

可知, 若能证明最后的二项接近于 0 就行了. 若对 $|x_t|$, 再次应用上面的方法, 于是就有

$$\frac{1}{2n}\int_{-(n+1)}^{n+1} |x_t|\mathrm{d}t = \frac{n+1}{n}\frac{1}{2(n+1)}\int_{-(n+1)}^{n+1} |x_t|\mathrm{d}t \to |x|^*, \quad \frac{1}{2n}\int_{-n}^{n} |x_t|\mathrm{d}t \to |x|^*,$$

因而它们的差

$$\frac{1}{2n}\int_{-(n+1)}^{-n} |x_t|\mathrm{d}t + \frac{1}{2n}\int_{n}^{n+1} |x_t|\mathrm{d}t \to 0,$$

故

$$\left|\frac{1}{2A}\int_{-A}^{-n} x_t \mathrm{d}t + \frac{1}{2A}\int_{n}^{A} x_t \mathrm{d}t\right| \leqslant \frac{1}{2n}\int_{-(n+1)}^{-n} |x_t|\mathrm{d}t + \frac{1}{2n}\int_{n}^{n+1} |x_t|\mathrm{d}t \to 0.$$

若令 $f(\xi_1,\xi_2,\cdots,\xi_n)$ 为 n 个变量的 Baire 函数, 则

$$y_t = f(x_{t_1+t}, x_{t_2+t}, \cdots, x_{t_n+t})$$

也为一可测的强平稳过程. 如果

$$E(|y_0|) = E(|f(x_{t_1}, x_{t_2}, \cdots, x_{t_n})|) < \infty,$$

则上述的定理可适用于 y_t , 因此极限

$$\lim_{A\to\infty} \frac{1}{2A}\int_{-A}^{A} f(x_{t_1+t}, x_{t_2+t}, \cdots, x_{t_n+t})\mathrm{d}t$$

也存在. 特别当 $E(|x_0|^2) < \infty$ 时, 由于 $E(|x_t\overline{x}_s|) < \infty$, 因此极限

$$\lim_{A\to\infty} \frac{1}{2A}\int_{-A}^{A} x_{t+\sigma}\overline{x}_{s+\sigma}\mathrm{d}\sigma$$

也存在. 由此可知下述极限的存在:

$$\begin{aligned} v^*(t,s) &= \lim_{A\to\infty} \frac{1}{2A}\int_{-A}^{A} (x_{t+\sigma} - x^*)\overline{(x_{s+\sigma} - x^*)}\mathrm{d}\sigma \\ &= \lim_{A\to\infty} \frac{1}{2A}\int_{-A}^{A} x_{t+\sigma}\bar{x}_{s+\sigma}\mathrm{d}\sigma - |x^*|^2. \end{aligned}$$

称 $V^* = (v^*(t,s))$ 为样本方差矩阵 (sample variance matrix).

由

$$v^*(t,s) = v^*(t-s)$$

及

$$\sum_{i,j} v^*(t_i - t_j)\xi_i\overline{\xi}_j = \lim_{A\to\infty} \frac{1}{2A}\int_{-A}^{A} \left|\sum \xi_i(x_{t_i+\sigma} - x^*)\right|^2 \mathrm{d}\sigma \geqslant 0,$$

可知 $v^*(t)$ 可以写成

$$v^*(t) = \int \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}S^*(\lambda) \quad (但S^*(-\infty) = 0).$$

$S^*(\lambda)$ 叫做样本谱函数 (sample spectral function). 由于

$$E(x^*) = 0 = E(x_0)$$

及

$$E(v^*(t)) = v(t) - E(|x^*|^2) \leqslant v(t),$$
$$E(S^*(\lambda)) = F(\lambda) - E(|x^*|^2)H(\lambda) \leqslant F(\lambda)$$
$$(H(\lambda) = 0(\lambda < 0), = 1(\lambda \geqslant 0)).$$

因此等号的成立必须 $E(|x^*|^2) = 0$, 即 $x^* = 0$. 由下式可知 $x^* = 0$ 与 $F(+0) = F(-0)$ 是等价的,

$$\begin{aligned} E(|x^*|^2) &= \lim_{A\to\infty}\left(\frac{1}{2A}\right)^2 \int_{-A}^{A}\int v(t-s)\mathrm{d}t\mathrm{d}s \\ &= F(+0) - F(-0). \end{aligned}$$

如前节所示一样,

$$x_t = \int \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}y_\lambda = \mathfrak{F}(\mathrm{D}y_\lambda),$$

而由此形式地计算 $v^*(t)$, 就得

$$\begin{aligned} v^*(t) &= \lim_{A\to\infty}\frac{1}{2A}\int_{-A}^{A}\iint \mathrm{e}^{-i2\pi[\lambda(t+s)-\mu s]}\mathrm{d}s\mathrm{d}y_\lambda\overline{\mathrm{d}y}_\mu \\ &= \iint \mathrm{e}^{-i2\pi\lambda t}\lim_{A\to\infty}\frac{1}{2A}\int_{-A}^{A}\mathrm{e}^{-i2\pi s(\lambda-\mu)}\mathrm{d}s\mathrm{d}y_\lambda\overline{\mathrm{d}y}_\mu \\ &= \iint \mathrm{e}^{-i2\pi\lambda t}\delta_{\lambda\mu}\mathrm{d}y_\lambda\overline{\mathrm{d}y}_\mu, \quad \delta_{\lambda\mu} = \begin{cases} 1 & (\lambda = \mu), \\ 0 & (\lambda \neq \mu), \end{cases} \\ &= \int \mathrm{e}^{-i2\pi\lambda t}|\mathrm{d}y_\lambda|^2. \end{aligned}$$

由此可以设想符号上的关系

$$\mathrm{d}S^*(\lambda) = |\mathrm{d}y_\lambda|^2.$$

这个有趣的事实可以利用 N. Wiener 的一般调和分析来严格地讨论. 上式的严密的意义就是: 对任意有界连续函数 $f(\lambda)$, 都成立

$$\int f(\lambda)\mathrm{d}S^*(\lambda) = \lim_{\varepsilon\downarrow 0}\int f(\lambda)\frac{|y_{\lambda+\varepsilon} - y_{\lambda-\varepsilon}|^2}{2\varepsilon}\mathrm{d}\lambda$$

§27 复正态系

我们称 $x_\alpha, a \in A$, 为一实正态系, 就是指随机向量 $\boldsymbol{x} = \prod_\alpha x_\alpha$ 的分布为 $\mathbb{R}^A(\boldsymbol{B}^A)$ 上的正态分布. Wiener 过程以及 (实) 正态平稳过程就是实正

态系. 上述的定义等价于: 对任意的 $\alpha_v \in A, z_v \in \mathbb{R}^1$,

$$E\left\{\exp\left(i\sum z_\nu x_{\alpha_\nu}\right)\right\} = \exp\left\{i\sum z_\nu m(\alpha_\nu) - \frac{1}{2}\sum z_\mu z_\nu v(\alpha_\mu, \alpha_\nu)\right\}$$

$$[m(\alpha) = E(x_\alpha), v(\alpha, \beta) = E\{(x_\alpha - m(\alpha))(x_\beta - m(\beta))\}].$$

特别当均值为 0 时, 就可以写成

$$E\{\exp(i\sum z_\nu x_{\alpha_\nu})\} = \exp\left\{-\frac{1}{2}\left\|\sum z_\nu x_{\alpha_\nu}\right\|^2\right\}$$

$$\left(\left\|x\right\|^2 = \int |x(\omega)|^2 P(\mathrm{d}\omega)\right).$$

将这个关系式推广到复随机变量系之后, 就可得 $x_\alpha, \alpha \in A$, 称它为复正态系, 如果对 $\alpha_\nu \in A, z_\nu \in \mathbb{C}(=\mathbb{R}^1 + i\mathbb{R}^1)$, 下式成立

$$E\{\exp(i\ \mathrm{Re}(\sum \overline{z}_\nu x_{\alpha_\nu}))\} = \exp\left\{-\frac{1}{4}\left\|\sum \overline{z}_\nu x_{\alpha_\nu}\right\|^2\right\}, \qquad (27.1)$$

(Re 表示实部; 横表示复共轭.)

特别地,

$$E\{\exp\{i\ \mathrm{Re}(\overline{z}x_\alpha)\}\} = \exp\left\{-\frac{|z|^2}{4}\|x_\alpha\|^2\right\},$$

这里, 若分别令 x_α 的实部和虚部为 x'_α 和 x''_α, 又设 $z = z' + iz''$, 则得

$$E\{\exp\{i(z'x'_\alpha + z''x''_\alpha)\}\} = \exp\left\{-\frac{\|x_\alpha\|^2}{4}(z'^2 + z''^2)\right\},$$

因而 x'_α 和 x''_α 是独立的, 且具有相同的正态分布 $N(\cdot; 0, \|x_\alpha\|^2/2)$.

定理 1 令 $x_\alpha, \alpha \in A$, 为复随机变量系, 又设 $Ex_\alpha = 0$. 若设

$$v(\alpha, \beta) = E(x_\alpha \overline{x}_\beta) = (x_\alpha, x_\beta),$$

则 $(v(\alpha, \beta))$ 是正定的, 也就是

$$\sum_{ij} \overline{\xi}_i \xi_j v(\alpha_i, \alpha_j) \geqslant 0. \qquad (27.2)$$

另一方面, 若 $(v(\alpha, \beta))$ 是正定的, 则存在复正态系 $x_\alpha, \alpha \in A$, 使得 (27.1) 成立.

证明 前一半:

$$\sum_{ij}\overline{\xi}_i\xi_j v(\alpha_i,\alpha_j)=\Big\|\sum_i\overline{\xi}_i x_{\alpha_i}\Big\|^2\geqslant 0.$$

后一半: 设 $\Omega=\mathbb{C}^A=\mathbb{R}^{2A}$, 又在 $\Omega(\boldsymbol{B}^{2A})$ 里引进如下的 (实) 正态分布 N : 在 $\mathbb{R}^{2A}$ 中的那些除在有限坐标以外, 其余的坐标为 0 的点的全体记为 $\mathbb{R}_0^{2A}$, 显然, 它与 $\mathbb{C}^A$ 中具有同一性质的点的全体 $\mathbb{C}_0^A$ 一致. 对 $\boldsymbol{z}(=\prod\limits_{\alpha}z_a)\in\mathbb{C}_0^A$, 设

$$\varphi(\boldsymbol{z})=\exp\Big\{-\frac{1}{4}\sum_{\alpha\beta}\overline{z}_\alpha z_\beta v(\alpha,\beta)\Big\}.$$

$\sum\limits_{\alpha\beta}$ 在外表上虽然是无限和, 但由于 $\boldsymbol{z}\in\mathbb{C}_0^A$, 因此只不过是有限和, 所以不会发生收敛的问题. 为了使得 $\varphi(\boldsymbol{z})$ 是 $\mathbb{R}^{2A}(\boldsymbol{B}^{2A})$ 上的某个正态分布 N 的特征函数, 只要证明: 把 $\sum\limits_{\alpha\beta}$ 看做 $\boldsymbol{z}\in\mathbb{R}_0^{2A}$ 的函数时, $\sum\limits_{\alpha\beta}$ 是正定的实二次形式. 这可由 $(v(\alpha,\beta))$ 的正定性得出. 事实上, 由正定性的假设得出 $v(\alpha,\beta)=\overline{v(\beta,\alpha)}$, 所以由此便知 $\sum\limits_{\alpha\beta}$ 是 $\boldsymbol{z}\in\mathbb{R}^{2A}$ 的实二次形式. 而它的正定性由 (27.2) 即可推出. 若令 $\Omega(\boldsymbol{B}^{2A},N)$ 为基本的概率空间, 又对 $\omega\in\Omega=\mathbb{C}^A$ 令 $x_\alpha(\omega)$ 为 ω 的 α 坐标 (复数), 则对 $\boldsymbol{z}(=\prod\limits_{\alpha}z_\alpha)\in\mathbb{C}_0^A$ 得

$$E\{\mathrm{e}^{i\mathrm{Re}\sum\limits_{\alpha}\overline{z}_\alpha x_\alpha}\}=E\{\mathrm{e}^{i\sum\limits_{\alpha}(z'_\alpha x'_\alpha+z''_\alpha x''_\alpha)}\}=\exp\Big\{-\frac{1}{4}\sum_{\alpha\beta}\overline{z}_\alpha z_\beta v(\alpha,\beta)\Big\}$$

$$(z_\alpha=z'_\alpha+iz''_\alpha,\ x_\alpha=x'_\alpha+ix''_\alpha).$$

以 $t\cdot z_\alpha$ 代替 z_α 便得

$$E\{\mathrm{e}^{it\ \mathrm{Re}\sum\limits_{\alpha}\overline{z}_\alpha x_\alpha}\}=\exp\Big\{-\frac{t^2}{4}\sum_{\alpha\beta}z_\alpha\overline{z}_\beta v(\alpha,\beta)\Big\}.$$

若两边对 t 微分两次, 然后再设 $t=0$, 则得

$$\frac{1}{4}\sum_{\alpha}\overline{z}_\alpha^2E(x_\alpha^2)+\frac{1}{4}\sum_{\alpha}z_\alpha^2E(\overline{x}_\alpha^2)+\frac{1}{2}\sum_{\alpha\beta}\overline{z}_\alpha z_\beta(x_\alpha,x_\beta)$$

$$=\sum_{\alpha\beta}\overline{z}_\alpha z_\beta v(\alpha,\beta).$$

因此

$$E(x_\alpha^2)=0,\quad (x_\alpha,x_\beta)=v(\alpha,\beta).$$

从而

$$E\{\mathrm{e}^{i\,\mathrm{Re}\sum\limits_\alpha \overline{z}_\alpha x_\alpha}\}=\exp\Big\{-\frac{1}{4}\Big\|\sum\overline{z}_\alpha x_\alpha\Big\|^2\Big\}.$$

注意 $E(x_\alpha^2)=0$ 是作为副产品而获得的, 但这是一个值得注意的性质. 事实上, 若设 $x_\alpha=x'_\alpha+ix''_\alpha$, 则如同前面已被注意到的那样, x'_α 和 x''_α 相互独立而且具有相同的分布 $N(\cdot;0,\|x_\alpha\|^2/2)$, 因而

$$E(x_\alpha^2)=E(x_\alpha'^2)-E(x_\alpha''^2)+2iE(x'_\alpha)E(x''_\alpha)=0.$$

定理 2 若 $x_\alpha,\alpha\in A$, 是复正态系, 并且是相互正交的, 则必为相互独立.

证明

$$\begin{aligned}
E(\mathrm{e}^{i\,\mathrm{Re}\sum\overline{z}_\alpha x_\alpha}) &=\exp\Big\{-\frac{1}{4}\Big\|\sum\overline{z}_\alpha x_\alpha\Big\|^2\Big\}\\
\text{由正交性}\qquad &=\exp\Big\{-\frac{1}{4}\sum|z_\alpha|^2\,\|x_\alpha\|^2\Big\}\\
&=\prod_\alpha\exp\Big\{-\frac{1}{4}|z_\alpha|^2\,\|x_\alpha\|^2\Big\}\\
&=\prod_\alpha E(\mathrm{e}^{i\,\mathrm{Re} z_\alpha x_\alpha}).
\end{aligned}$$

由此可知 $\boldsymbol{x}=\prod\limits_\alpha x_\alpha$ 的分布是各个 x_α 的分布的直积.

定理 3 令 $x_\alpha(\alpha\in A)$ 为复正态系, 则 $\mathrm{Re}(x_\alpha)$ $(\alpha\in A)$ 和 $\mathrm{Im}(x_\alpha)$, $(\alpha\in A$, Im 是虚部) 都是实正态系.

只需在 (27.1) 中, 设 $\mathrm{Im}(z_\nu)$ 或者 $\mathrm{Re}(z_\nu)=0$ 即可得证. 又由定义不难证明下述的定理.

定理 4 若令 $x_\alpha(\alpha\in A)$ 和 $y_\alpha,(\alpha\in A)$ 都是实正态系, 又令两系相互独立, 则 $x_\alpha+iy_\alpha(\alpha\in A)$ 是复正态系.

定理 5 若 $x_\alpha(\alpha\in A)$ 是复正态系, 且设每一 $y_\beta(\beta\in B)$ 是 $x_\alpha(\alpha\in A)$ 的复系数的线性组合或者是这种线性组合依范数的极限, 则 $y_\beta(\beta\in B)$ 也是正态系.

例 复 Wiener 过程(complex Wiener process) 若 $x_t(-\infty<t<\infty)$ 是复正态系, 而且具有正交增量, 即

$$\text{对} t < s \leqslant u < v, \quad (x_s - x_t, x_v - x_u) = 0,$$

则 $x_t(-\infty < t < \infty)$, 就叫做复 Wiener 过程. 这时根据定理 5 可知, 对于 $t_1 < s_1 \leqslant t_2 < s_2 \leqslant \cdots \leqslant t_n < s_n, x_{s_i} - x_{t_i}(i = 1, 2, \cdots, n)$ 也是复 Wiener 过程, 而且是相互正交的, 从而由定理 2 可知是独立的. 因而它是一个复可加过程. $x_s - x_t$ 的分布对于环绕着复平面上的原点的旋转是不变的分布. 因为 $s < t < u$ 时 $(x_u - x_t, x_t - x_s) = 0$, 所以

$$\| x_u - x_s \|^2 = \| x_u - x_t \|^2 + \| x_t - x_s \|^2,$$

而且除去一附加常数以外, 可以唯一决定一个增函数 $F(t)$, 使得

$$\| x_t - x_s \|^2 = F(t) - F(s).$$

特别地, 若 $F(t)$ 是右连续的, 则 x_t 也在范数的意义下为右连续的. 我们还可以证明下述的逆命题: 对任一给定的增函数 $F(t)$, 必存在一复 Wiener 过程 x_t, 使得 $\| x_t - x_s \|^2 = F(t) - F(s)$. 证明如下. 以 A 表示实数区间 $(a, b](-\infty < a < b < \infty)$ 的全体. 若对 $\alpha, \beta \in A$, 定义

$$v(\alpha, \beta) = \begin{cases} F\text{在区间}\alpha \cap \beta\text{上的增量} & (\alpha \cap \beta \neq \varnothing\text{即}\alpha \cap \beta = \text{区间}), \\ 0 & (\alpha \cap \beta = \varnothing), \end{cases}$$

则 $(v(\alpha, \beta))$ 是正定的. 事实上, 若令 α 的示性函数为 $c(t, \alpha)$, 则

$$\begin{aligned} \sum_{ij} \overline{\xi}_i \xi_j v(\alpha_i, \alpha_j) &= \sum_{ij} \overline{\xi}_i \xi_j \int c(t, \alpha_i) c(t, \alpha_j) \mathrm{d}F(t) \quad (\text{Riemann-Stieltjes 积分}) \\ &= \int \Big| \sum_i c(t, \alpha_i) \overline{\xi}_i \Big|^2 \mathrm{d}F(t) \geqslant 0. \end{aligned}$$

故存在复正态系 $\{x_\alpha\}$, 使得

$$(x_\alpha, x_\beta) = v(\alpha, \beta).$$

若 $t < s < u$, 则

$$\begin{aligned} &\| x_{(t,s]} + x_{(s,u]} - x_{(t,u]} \|^2 \\ &= F(s) - F(t) + F(u) - F(s) + F(u) - F(t) - 2(F(s) - F(t)) - 2(F(u) - F(s)) \\ &= 0, \end{aligned}$$

因而

$$x_{(t,s]}+x_{(s,u]}=x_{(t,u]}(\text{a.e.}).$$

故若令 $\alpha<\min(0,t)$, 且设

$$x_t=x_{(a,t]}-x_{(a,0]},$$

则 x_t 与 α 的选择无关 (除概率为 0 以外), 于是 $x_t(-\infty<t<\infty)$ 就是所求的复 Wiener 过程.

特别在 $F(t)=t$ 时, 即过程对时间是齐次的, Wiener 把这种过程叫做**Brown 运动**.

§28 正态平稳过程

复正态系 $x_t(-\infty<t<\infty)$ 是弱平稳过程时, 就称它为复正态弱平稳过程. 对应于 §22 的定理 2, 就有下述定理.

定理 1 复正态弱平稳过程是强平稳的.

证明 令 $x_t(-\infty<t<\infty)$ 为复正态弱平稳过程. 对任意的 $t_1<t_2<\cdots<t_n$ 和 h,

$$\begin{aligned}E\{\mathrm{e}^{i\,\mathrm{Re}(\sum\overline{z}_\nu x_{t_\nu})}\}&=\exp\left\{-\frac{1}{4}\sum\overline{z}_\mu z_\nu(x_{t_\mu},x_{t_\nu})\right\}\\&=\exp\left\{-\frac{1}{4}\sum\overline{z}_\mu z_\nu v(t_\mu-t_\nu)\right\},\quad (x_t,x_s)=v(t-s)\\&=E\{\mathrm{e}^{i\,\mathrm{Re}(\sum\overline{z}_\nu x_{t_\nu+h})}\}.\end{aligned}$$

故 (x_{t_ν}) 的分布与 $(x_{t_\nu+h})$ 的分布相等, 于是 (x_t) 就成为强平稳的.

由这定理可知, 只提**复正态平稳过程**就够了. 又因为通常我们总是假定平稳过程是复的, 所以简称它为**正态平稳过程**.

在前面已经证明: 由 Khinchin 定理, $v(t)=(x_{s+t},x_s)$ 的谱分解是可能的. 但没有指明对这样的函数 $v(t)$, 是否存在平稳过程 (x_t), 使得 $(x_t,x_s)=v(t-s)$. 事实上, 这是存在的, 甚至存在满足这个条件的正态平稳过程.

定理 2 若令 $v(t)=\displaystyle\int\mathrm{e}^{-i2\pi\lambda t}\mathrm{d}F(\lambda)$, $F(\lambda)$ 是有界右连续的增函数, 则存在正态平稳过程 $x_t(-\infty<t<\infty)$, 使得

$$(x_t, x_s) = v(t-s).$$

证明 因对任意的 t_μ, ξ_μ,

$$\sum \bar{\xi}_\mu \xi_\nu v(t_\mu - t_\nu) = \int \left| \sum \bar{\xi}_\mu \mathrm{e}^{-i2\pi t_\mu \lambda} \right|^2 \mathrm{d}F(\lambda) \geqslant 0,$$

所以由前节定理 1 立即知道 (x_t) 的存在.

其次考虑 x_t 的 Kolmogoroff 谱分解:

$$x_t = \int \mathrm{e}^{-i2\pi t\lambda} \mathrm{d}y_\lambda$$

由于 y_λ 属于由 x_t 所张成的 Hilbert 穿间 $\boldsymbol{H}_0$, 所以由前节定理 5 便知 $y_\lambda(-\infty < \lambda < \infty)$ 也是复正态系. 又因 $y_\lambda(-\infty < \lambda < \infty)$ 具有正交增量, 所以是复 Wiener 过程. 利用 Khinchin 谱函数 F, 就得

$$\| y_\lambda - y_\mu \|^2 = F(\lambda) - F(\mu) \qquad (\lambda > \mu).$$

y_λ 不一定是齐次的. 故得以下结论.

定理 3 正态平稳过程就是复 Wiener 过程的微分 (广义函数的意义下) 的 Fourier 变换.

§29 Wiener 积分, 多重 Wiener 积分

令 $x_t(-\infty < t < \infty)$ 为一复 Wiener 过程, 又设

$$\| x_t - x_s \|^2 = F(t) - F(s) \qquad (t > s).$$

因为 x_t 具有正交增量, 所以由 §23 可知, 此时可定义形状为

$$I(f) = \int f(t) \mathrm{d}x_t, \quad f \in L^2(\mathbb{R}^1, \mathrm{d}F)$$

的积分. 若 (x_t) 是复 Wiener 过程, 则特别称此积分为 **Wiener 积分**. 其次, 在这种场合下, 虽可定义下述形状的**多重 Wiener 积分**:

$$I_{p,q}(f) = \int \cdots \int f(t_1, \cdots, t_p, s_1, \cdots, s_q) \mathrm{d}x_{t_1} \cdots \mathrm{d}x_{t_p} \mathrm{d}\overline{x}_{s_1} \cdots \mathrm{d}\overline{x}_{s_q},$$

但为此必须假定 $F(t)$ 是连续的. 此处不来讨论它的详细证明, 而只就 $p=2, q=0$ 的场合

$$I(f)=I_{2,0}(f)=\iint f(t_1,t_2)\mathrm{d}x_{t_1}\mathrm{d}x_{t_2}$$

来说明一下证明的要点. 首先, 在 f 是与对角线无公共点的二维区间 $(a,b\,]\times(c,d\,]$ 的示性函数时, 定义

$$I(f)=(x_b-x_a)(x_d-x_c).$$

这里, 与对角线无公共点的条件 (姑且叫做**条件A**) 是重要的. 若 f 是上述那样的示性函数的线性组合 (记其全体为 S), 则定义它为上面那种 $I(f)$ 的线性组合. 显然 I 是从 S 到 $\boldsymbol{H}=L^2(\Omega)$ 中的线性算子. 利用条件 A 就得

$$\|I(f)\|^2\leqslant\|f\|^2\equiv\iint|f(t_1,t_2)|^2\mathrm{d}F(t_1)\mathrm{d}F(t_2).$$

特别地, 若 $f(t_1,t_2)$ 关于 (t_1,t_2) 对称, 则上式的不等号变为等号, 又若 f,g 都对称, 则得

$$(If,Ig)=(f,g).$$

又由 F 的连续性, 得 $\overline{S}=L^2\equiv L^2(\mathbb{R}^2,(\mathrm{d}F)^2)$, 因而可以把 $I(f)$ 推广到 L^2 上去.

对于一般的 $I_{p,q}(f)$ 也可以同样定义. 令 $I_{p,q}(f)$ 的象的全体为 $\boldsymbol{H}_{p,q}$. 特别令 $\boldsymbol{H}_{00}$ 为 $L^2(\Omega)$ 中的常数的全体所形成的一维空间. 由于条件 A 的作用, 于是 $\{\boldsymbol{H}_{p,q}\}_{p,q}$ 就成为相互正交的 $\boldsymbol{H}$ 的子空间. 其次, 若令 $\boldsymbol{H}^*$ 为关于 $x_t(-\infty<t<\infty)$ 可测而且 $E(|x|^2)<\infty$ 的复数值 (随机变量)x 的全体, 则可以证明 $\boldsymbol{H}^*$ 是 $\boldsymbol{H}_{p,q}$ 的直和. 因此, 可将 $x\in\boldsymbol{H}^*$ 展为直交和

$$x=\sum_{p,q}I_{p,q}(f_{p,q}).$$

故可将 $f_{p,q}(t_1,\cdots,t_p,s_1,\cdots,s_q)$ 取为分别关于 $(t_1,\cdots,t_p)(s_1,\cdots,s_q)$ 的各组为对称的函数. 此时, 这些函数可由 x 唯一地确定.

例 1 若令 $x_t(-\infty<t<\infty)$ 为对时间齐次的复 Wiener 过程, 又设

$$y_t=\int f(t+s)\mathrm{d}x_s,\quad f\in L^2(\mathbb{R}^1),$$

则 y_t 是正态平稳过程. 事实上,

$$(y_t,y_u)=\int f(t+s)\overline{f(u+s)}\mathrm{d}s=\int f(t-u+s)\overline{f(s)}\mathrm{d}s,$$

因而 (y_t, y_u) 是 $t-u$ 的函数, 于是得出平稳性. 至于正态性, 则由 y_t 是 x_s 的线性组合的极限这一事实即得到说明. 现在若令 f 的逆 Fourier 变换为 $\widehat{f}$, 则得

$$(y_t, y_u) = \int \mathrm{e}^{-i2\pi(t-u)\lambda} |\widehat{f}(\lambda)|^2 \mathrm{d}\lambda.$$

故 $|\widehat{f}(\lambda)|^2$ 是谱函数的微分系数.

例 2 依照与上述一样的方法, 利用多重 Wiener 积分, 若令

$$z_t = \iint f_2(t_1+t, t_2+t)\mathrm{d}x_{t_1}\mathrm{d}x_{t_2} + \int f_1(t_1+t)\mathrm{d}x_{t_1},$$

则 z_t 是一强平稳过程, 但除 $f_2 \equiv 0$ 的情形以外, 并不具有正态性. 进一步利用更高维数的多重 Wiener 积分以及它的极限, 所得到的仍然是强平稳过程. 但考察一下是否能利用这种方法来构成一般的强平稳过程, 这将是一个有意义的问题.

§30 正态平稳过程的遍历性

首先来给出强平稳过程 $x_t(-\infty < t < \infty)$ 的遍历性和强混合性的定义. 若令 x 是关于 $x_t(-\infty < t < \infty)$ 可测的随机变量, 则可以写成

$$f_n(x_{t_1}, \cdots, x_{t_n}) \to x(\text{依概率收敛}). \tag{30.1}$$

此处 f_n 是 n 个复变量的 Baire 函数. 由强平稳性得

$$\begin{aligned} &P\{\omega/|f_n(x_{t_1+t}, \cdots, x_{t_n+t}) - f_m(x_{t_1+t}, \cdots, x_{t_m+t})| > \varepsilon\} \\ =&P\{\omega/|f_n(x_{t_1}, \cdots, x_{t_n}) - f_m(x_{t_1}, \cdots, x_{t_m})| > \varepsilon\} \\ &\to 0 \quad (m, n \to \infty), \end{aligned}$$

所以 $f_n(x_{t_1+t}, \cdots, x_{t_n+t}), n = 1, 2, \cdots$, 依概率收敛于某个随机变量 x'. x' 以概率为 1 地仅由 x 和 t 决定, 并且与 (30.1) 的序列的选择无关. 记 $x' = T_t x$, 于是就有

$$T_{t+s}x = T_t T_s x \text{ (a.e.)}. \tag{30.2}$$

若对所有的 $t, T_t x = x$ 时, 则 x 叫做不变的. $x \equiv$ 常数就是不变的, 若除此以外不再存在不变的 x 时, $x_t(-\infty < t < \infty)$ 就叫做**具有遍历**

性(ergodic). 若有一不变的 x, 考虑其截口 $x^{(M)}[x^{(M)}(\omega)=x(\omega)(|x(\omega)|\leqslant M),=0(|x(\omega)>M)]$, 它显然也是不变的. 若 x 不是常数, 则可取 M 足够大, 使得 $x^{(M)}$ 也不是常数, 因而若除去常数外不存在有界不变的 x, 则过程就具有遍历性. 其次, 对于依范数有界且关于 $x_t(-\infty<t<\infty)$, 可测的任意 x 与 y, 都有

$$E(T_t x\cdot\bar{y})\to E(x)E(\bar{y})\quad 即(T_t x,y)\to(x,1)(1,y)$$

时, 则称 $x_t(-\infty<t<\infty)$ 为具有**强混合性**(strongly mixing). 强混合性的条件要比遍历性来得强. 其原因是: 若令 $T_t x=x$, 则由强混合性得

$$E(x^2)=E(T_t x\cdot x)\to E(x)^2,$$

因此 $V(x)=E(x^2)-E(x)^2=0$, 于是 $x=$ 常数.

有了上面这些准备, 可以来证明下面的定理.

定理 1(G. Maruyama) 欲使正态平稳过程具有遍历性, 它的必要而且充分的条件是谱函数 $F(\lambda)$ 为连续.

定理 2 欲使正态平稳过程具有强混合性, 它的必要而且充分的条件是 $v(t)\to 0(t\to\infty)$.

令 $x_t(-\infty<t<\infty)$ 为正态平稳过程, 且令 $F(\lambda)$ 在 $\lambda=\mu$ 处有跳跃. 令 x_t 的 Kolmogoroff 谱分解为

$$x_t=\int \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}y_\lambda,$$

如果沿用证明前一定理时所采用的 $U_t,E(\lambda)$ 等记号, 那么 $y_\lambda=E(\lambda)\cdot x_0$ 也在 $\lambda=\mu$ 处有跳跃, 而且它的跃度 $z=y_{\mu+0}-y_{\mu-0}$ 满足 $U_t z=\mathrm{e}^{-i2\pi\mu t}z$. 因为 T_t 是 U_t 的推广, 所以 $T_t z=\mathrm{e}^{-i2\pi\mu t}z$. 由 T_t 的定义可得 $|T_t z|=T_t|z|$, 故 $T_t|z|=|z|$. 即 $|z|$ 是不变的. 又因 z 的分布是在复平面上旋转不变的正态分布, 且由于 $\|z\|^2=F(\mu+0)-F(\mu-0)>0$, 所以不是 δ 分布, 从而 $|z|$ 不是常数. 由此就证明了定理 1 中的必要性. 其次往证充分性. 令 x 为有界不变. 因为关于 $y_\lambda(-\infty<\lambda<\infty)$ 可测的全体和关于 $x_t(-\infty<t<\infty)$ 可测的全体是一致的, 所以 x 关于 $y_\lambda(-\infty<\lambda<\infty)$ 是可测的, 又因 $\|x\|^2<\infty$, 所以由前节的结果而得

$$x=常数+\sum_{p+q>0}\int\cdots\int f_{pq}(\lambda_1\cdots\lambda_p,\mu_1\cdots\mu_q)\\ \mathrm{d}y_{\lambda_1}\cdots\mathrm{d}y_{\lambda_p}\mathrm{d}\bar{y}_{\mu_1}\cdots\mathrm{d}\bar{y}_{\mu_q}.\tag{30.3}$$

假设 $f_{p,q}$ 分别关于 $(\lambda_i)(\mu_j)$ 对称. 为了求 $T_t x$, 只要令 $\mathrm{d}y_{\lambda_1}\cdots$ 为 $T_t\mathrm{d}y_{\lambda_1}\cdots$ 就可以了, 又因 $T_t\mathrm{d}y_\lambda = \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}y_\lambda, T_t\mathrm{d}\overline{y}_\mu = \mathrm{e}^{i2\pi\mu t}\mathrm{d}\overline{y}_\mu$ (这是形式上的说法. 严密地说, 就要稍长一些, 但由此不难充分领会它的意思), 所以

$$T_t x = 常数 + \sum_{p+q>0}\int\cdots\int f_{pq}(\lambda_1\cdots\lambda_p,\mu_1\cdots\mu_q) \cdot\mathrm{e}^{-i2\pi t(\sum\lambda_\pi-\sum\mu_\rho)}\mathrm{d}y_{\lambda_1}\cdots\mathrm{d}\overline{y}_{\mu_q}. \tag{30.4}$$

因被积函数分别关于 $(\lambda_x)(\mu_\rho)$ 对称, 所以由 $T_t x = x$ 可得

$$f_{pq}(\lambda_1\cdots\lambda_p,\mu_1\cdots\mu_q)\mathrm{e}^{-i2\pi t(\sum\lambda_\pi-\sum\mu_\rho)} = f_{pq}(\lambda_1\cdots\lambda_p,\mu_1\cdots\mu_q).$$

故在超平面 $\Pi:\sum\lambda_\pi-\sum\mu_\rho=0$ 之外 $f_{pq}=0$. 又由 F 的连续性得

$$\int_\Pi \mathrm{d}F(\lambda_1)\cdots\mathrm{d}F(\lambda_p)\mathrm{d}F(\mu_1)\cdots\mathrm{d}F(\mu_q)=0,$$

所以 $f_{pq}=0$(a.e.). 故得出 $x=$ 常数.

现在转到定理 2 的证明上去. 若具有强混合性, 则 $v(t)=(x_t,x_0)=(T_t x_0,x_0)\to(x_0,1)^2=0$. 其次, 假定 $v(t)\to 0$. 由此可知 $F(\lambda)$ 必连续. 故关于 $x_t(-\infty<t<\infty)$ 可测且依范数有界的 x,y 可写成 (30.3) 的形状, 再利用前节所述 $\boldsymbol{H}_{pq}$ 的正交性就得

$$(T_t x,y)=(x,1)(1,y)+\sum_{p+q>0}\int\cdots\int f_{pq}(\cdots)\overline{g}_{pq}(\cdots) \cdot\mathrm{e}^{-i2\pi t(\sum\lambda_\pi-\sum\mu_\rho)}\mathrm{d}F(\lambda_1)\cdots\mathrm{d}F(\mu_q). \tag{30.5}$$

再由 $v(t)\to 0$ 可得, 对于任意的 $-\infty<a<b<\infty$,

$$\left|\int_a^b \mathrm{e}^{-i2\pi t\lambda}\mathrm{d}F(\lambda)\right|\to 0\quad (t\to\infty).$$

若以区间的特征函数的线性组合来逼近 f_{pq} 和 g_{pq}, 则 (30.5) 中的 $\sum$ 内的各项收敛于 0. 故若 x,y 的展式只有有限项时, 就得 $(T_t x,y)\to(x,1)(1,y)$. 用这些 x,y 依范数逼近一般的 x,y, 就可以证明同样的结论. 这时必须注意到 $\| T_t x \|=\| x \|$.

例　因为正态平稳过程 x_t 是强平稳过程, 又因 $E(|x_0|)<\infty$, 所以样本均值:

$$x^* = \lim_{A\to\infty} \frac{1}{2A} \int_{-A}^{A} x_t \mathrm{d}t$$

存在. 显然 $T_t x^* = x^*$, 所以若 $F(\lambda)$ 为连续, 则由定理 1 得出 x^* =常数. 故 $x^* = E(x^*) = E(x_0)$. 这就是用样本过程来求 $E(x_0)$ 的公式. 同样的事情也可以适用于 $v(t)$.

§31 平稳过程的普遍化

首先叙述**平稳序列**. 这虽然不应叫做普遍化, 但顺便叙述一下. 关于随机序列 $x_n(n = \cdots, -2, -1, 0, 1, 2, \cdots)$ 的弱平稳性、强平稳性和正态性的定义与平稳过程的场合相同. 又在正态平稳序列里没有必要区别强平稳和弱平稳, 这也与以前一样. 若假定 $E(x_n) = 0$, 则 $v(n) = (x_{n+m}, x_m)$ 的 Khinchin 谱分解为

$$v(n) = \int_0^1 \mathrm{e}^{-i2\pi\lambda n} \mathrm{d}F(\lambda), \tag{31.1}$$

而 $\mathrm{d}F$ 就是 $\mathbb{R}^1/\mathrm{mod}\ 1$ 上的测度. 又 x_n 的 Kolmogoroff 谱分解为

$$x_n = \int_0^1 \mathrm{e}^{-i2\pi\lambda n} \mathrm{d}y_\lambda, \quad \| \mathrm{d}y_\lambda \|^2 = \mathrm{d}F(\lambda). \tag{31.2}$$

由此可知, 平稳过程的性质几乎全部保持, 而且以更简单的形态出现于平稳序列.

平稳广义过程(stationary random distribution) 随机过程 $x_t(\omega)$ 可以考虑为以 ω 作为辅助变量的 t 的函数, 但是如果所考虑的以 ω 作为辅助变量的 t 的函数为广义函数, 则就可以考虑所谓**广义随机过程**(random distribution). 广义随机过程也可以分为强弱两种来考虑. 令 $\mathfrak{D}$ 为无限可微而且支集为有界的 t 的函数的全体, 再令广义函数的全体为 $\mathfrak{D}'$. 若有 $\phi \in \mathfrak{D}$ 和 ω 的函数 $x(\phi, \omega)$, 使得对 ω 的所有的值 (或者几乎所有的值), 把 $x(\phi, \omega)$ 看做 ϕ 的函数时它属于 $\mathfrak{D}'$, 且对任意给定的 $\phi, x(\phi, \omega)$ 是 ω 的可测函数时, $x(\phi, \omega)$ 就叫做**强义的广义过程**. 反之, 若对任意给定的 ϕ, 把 $x(\phi, \omega)$ 看做 ω 的函数而属于 $\boldsymbol{H} = L^2(\Omega)$, 且映照

$$x : \mathfrak{D} \ni \phi \to x(\phi, \cdot) \in \boldsymbol{H}$$

为取值于 $\boldsymbol{H}$ 的广义函数时, 即 $x \in \mathfrak{D}'_{\boldsymbol{H}}$ 时, $x(\phi,\omega)$ 就叫做**弱义的广义过程**. 特别地, 若强义的广义过程满足**强义的平稳性**:

$$\text{“}(x(\phi_1),\cdots,x(\phi_n))\ \text{的分布和}\ (x(\phi_1^{(h)}),\cdots,x(\phi_n^{(h)}))\ \text{的分布恒相同”}$$

$$(\phi^{(h)}(t)=\phi(t+h)), \tag{31.3}$$

则叫做**强平稳广义过程**. 设 $x_t(\omega)$ 为一可测的强平稳过程, 且设 $E(|x_0|)<\infty$. 对此, 若定义

$$x(\phi,\omega)=\int\phi(t)x_t(\omega)\mathrm{d}t, \tag{31.4}$$

则这就是强平稳广义过程. 事实上, 由于 $E(|x_0|)<\infty$, 便知对几乎所有的 ω,

$$\int\frac{|x_t(\omega)|}{1+t^2}\mathrm{d}t<\infty \tag{31.5}$$

(试考虑积分的平均值), 由此即知, 作为 t 的函数, $x_t(\omega)$ 是局部可积的. 因此 (31.4) 给出了强义的广义过程. 还可证明它具有强义的平稳性 (31.3). 又若有一弱义的广义过程 $x(\phi,\omega),\phi\in\mathfrak{D}$, 且

$$m(\phi)=E(x(\phi)),\quad v(\phi,\psi)=E((x(\phi)-m(\phi))\overline{(x(\psi)-m(\psi))}) \tag{31.6}$$

对平行移动都为不变:

$$m(\phi^{(h)})=m(\phi),\quad v(\phi^{(h)},\psi^{(h)})=v(\phi,\psi), \tag{31.7}$$

则就叫做**弱平稳广义过程**. 例如若 x_t 为一弱平稳过程, 则下式所定义的就可以看做是弱平稳广义过程:

$$x(\phi)=\int\phi(t)x_t\mathrm{d}t\quad(\text{积分是对依范数收敛而言的}).$$

事实上,

$$\begin{aligned}
m(\phi)&=\int\phi(t)m(t)\mathrm{d}t=m\int\phi(t)\mathrm{d}t,\\
v(\phi,\psi)&=\iint\phi(t)\overline{\psi(s)}v(t,s)\mathrm{d}t\mathrm{d}s=\iint\phi(t)\overline{\psi(s)}v(t-s)\mathrm{d}t\mathrm{d}s\\
&=\int v(t)\int\phi(t+s)\overline{\psi(s)}\mathrm{d}s\mathrm{d}t\\
&=\int v(t)\int\phi(t-s)\check{\psi}(s)\mathrm{d}s\mathrm{d}t,\quad\check{\psi}(s)=\overline{\psi(-s)},
\end{aligned}$$

$$= \int v(t)(\phi * \check{\psi})(t)\mathrm{d}t$$
$$= v(\phi * \check{\psi}).$$

由此就得出 m, v 对平行移动的不变性, 又知道 $v(\phi,\psi)$ 可以写成 $v(\phi * \psi)$ 的形状. 因为后面的事实可由 $v(\phi,\psi)$ 是有对平行移动的不变性而推出, 所以对一般的弱平稳广义过程, 恒有

$$v(\phi,\psi) = v(\phi * \check{\psi}). \tag{31.8}$$

由于这里的 v 可以看做 $\mathfrak{D}'$ 的元素, 又因

$$v(\phi * \check{\phi}) = v(\phi,\phi) = \| x(\phi) - m(\phi) \|^2 \geqslant 0,$$

所以利用推广到广义函数情形的 Bochner 定理[①], 就得

$$v = \mathfrak{D}'\text{-} \lim_{A\to\infty} \int_{-A}^{A} \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}F(\lambda), \tag{31.9}$$

此处

$$\mathrm{d}F(\lambda) \geqslant 0, \quad \int \frac{\mathrm{d}F(\lambda)}{(1+\lambda^2)^k} < \infty. \tag{31.10}$$

这就是 A. Khinchin 谱分解的普遍化. 同样, 对于 $x(\phi)$, 可找到正交增量过程 y_λ, 使得

$$x = \mathfrak{D}'_{\boldsymbol{H}}\text{-} \lim_{A\to\infty} \int_{-A}^{A} \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}y_\lambda, \quad \| \mathrm{d}y_\lambda \|^2 = \mathrm{d}F(\lambda). \tag{31.11}$$

这就是 Kolmogoroff 谱分解的普遍化.

现在令 $(x_t, -\infty < t < \infty)$ 为对时间齐次的复 Wiener 过程 (§27 末尾). x_t 本身显然不是平稳过程, 然而由于对时间的齐次性 $\| \mathrm{d}x_t \|^2 = \mathrm{d}t$, 所以若考虑 x_t 在广义函数 $(\mathfrak{D}'_{\boldsymbol{H}})$ 的意义下的微分 (由 $\| \mathrm{d}x_t \|^2 = \mathrm{d}t$ 而明显地可看出普通意义下的微分是不存在的), 且记为 $\mathrm{D}x_t$ 时, 则这就成为弱平稳广义过程. 事实上, 由

① 参见 Schwartz, Théorie des distributions, II. Paris, Herman, 1952,432.

—— 俄译本注

$$\mathrm{D}x_t(\phi) = -x_t(\phi') = -\int x_t\phi'(t)\mathrm{d}t = -\int (x_t - x_a)\phi'(t)\mathrm{d}t,$$

(a 为任意) 所以 $m(\phi)$ 与 $v(\phi,\psi)$ 可以写成

$$\begin{aligned}
m(\phi) &= 0,\\
v(\phi,\psi) &= E\left\{\iint \phi'(t)\overline{\psi'(s)}(x_t - x_a)\overline{(x_s - x_a)}\mathrm{d}t\mathrm{d}s\right\},\\
v(\phi,\psi) &= \int_a\int_a \phi'(t)\overline{\psi'(s)}(x_t - x_a, x_s - x_a)\mathrm{d}t\mathrm{d}s\\
&\quad [\text{取 } a \text{ 使得 } \phi,\psi \text{ 的支集落入}(a,\infty)]\\
&= \int_a\int_a \phi'(t)\overline{\psi'(s)}(\min(t,s) - a)\mathrm{d}t\mathrm{d}s\\
&= \int \phi(t)\overline{\psi(t)}\mathrm{d}t = (\phi * \check{\psi})(0)\\
&= \delta(\phi * \check{\psi}) \quad (\delta\text{是 Dirac 的 } \delta \text{ 函数}).
\end{aligned}$$

这就意味着 v 对平行移动是不变的, 又若写成 $v(\phi,\psi) = v(\phi * \check{\psi})$, 则得 $v = \delta$. 若引用下述记号

$$\begin{aligned}
\left(\frac{\mathrm{d}x_t}{\mathrm{d}t}, \overline{\frac{\mathrm{d}x_s}{\mathrm{d}s}}\right) &= 0 \quad (t \neq s),\\
\left(\frac{\mathrm{d}x_t}{\mathrm{d}t}, \overline{\frac{\mathrm{d}x_t}{\mathrm{d}t}}\right) &= \frac{\|\mathrm{d}x_t\|^2}{\mathrm{d}t^2} = \frac{\mathrm{d}t}{\mathrm{d}t^2} = \frac{1}{\mathrm{d}t},
\end{aligned}$$

则可使我们更易领会这一点. 又因

$$\delta = \mathfrak{D}'\text{-}\lim_{A\to\infty}\int_{-A}^{A} \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}\lambda,$$

所以 $F(\lambda) \equiv \lambda$ 就是谱函数. 这时因为

$$\int \frac{\mathrm{d}\lambda}{1+\lambda^2} < \infty,$$

所以 (31.10) 中的指数 k 就是 1.

向量值平稳过程 可以考虑使得 $\boldsymbol{x}_t(-\infty < t < \infty)$ 的值为 m 维向量的平稳过程. 今就弱平稳过程作一些简单的说明. 设 $\boldsymbol{m}(t) \equiv E(\boldsymbol{x}_t) = 0$, 这并不丧失普遍性. 令 $\boldsymbol{v}(t,s)$ 是矩阵 $(v_{ij}(t,s))$, 而

$$v_{ij}(t,s) = E(x_t^i x_s^j) \quad \text{即} \quad \boldsymbol{v}(t,s) = E(\boldsymbol{x}_t \times \boldsymbol{x}_s). \tag{31.12}$$

当 $v_{ij}(t,s)$ 仅是 $t-s$ 的函数时, $\boldsymbol{x}_t$ 就叫做弱平稳的, 记为 $\boldsymbol{v}_{ij}(t,s)=v_{ij}(t-s)$. 至于 $v(t)=(v_{ij}(t))$ 的谱分解, H. Cramér 已经采用 A. Khinchin 关于谱分解的普遍化的形状求出了. 那就是有一取值于 $m\times m$ 阶矩阵的 $\boldsymbol{F}(\lambda)=(F_{ij}(\lambda))$, 使得

$$\boldsymbol{v}(t)=\int \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}\boldsymbol{F}(\lambda). \tag{31.13}$$

此处 $\boldsymbol{F}(\lambda)$ 是满足下列关系式的 Hermite 矩阵:

$$\lambda\geqslant\mu\Rightarrow\boldsymbol{F}(\lambda)-\boldsymbol{F}(\mu)\geqslant 0\quad(\geqslant 0\text{是正定型}),$$
$$\boldsymbol{F}(\infty)\text{确定},\ \boldsymbol{F}(-\infty)=0,$$
$$\boldsymbol{F}(\lambda+0)=\boldsymbol{F}(\lambda),$$

事实上, 因对任意的 $\{a_i\}, v_a(t)=\sum a_i\bar{a}_j v_{ij}(t)$ 满足

$$\sum_{\mu\nu}\xi_\mu\bar{\xi}_\nu v_a(t_\mu-t_\nu)=\sum_{\mu\nu}\sum_{ij}\xi_\mu\bar{\xi}_\nu a_i\bar{a}_j(x^i_{t_\mu},x^j_{t_\nu})=\left\|\sum\xi_\mu a_i x^i_{t_\mu}\right\|^2\geqslant 0,$$

所以
$$v_a(t)=\int \mathrm{e}^{-i2\pi\lambda t}\mathrm{d}F_a(\lambda).$$

$\Delta F_a(\lambda)$ 就是 (a_i) 的 Hermite 形式而且 $\Delta F_a(\lambda)\geqslant 0$. 因此可以确定 $F_{ij}(\lambda)$, 使得 $\Delta F_a(\lambda)=\sum a_i\bar{a}_j\Delta F_{ij}(\lambda)$, 而 $\Delta F(\lambda)=(\Delta F_{ij}(\lambda))$ 就是正定的.

在这个场合下也可以得到相当于 Kolmogoroff 分解的谱分解.

关系到时间和地点的随机过程 随机过程是表达随时间而变化的偶然量, 但也可以考虑时间和地点双方都改变的情形. 譬如说, 在某个时刻 t 处于某一地点 $\boldsymbol{\xi}=(\xi_1,\xi_2,\xi_3)$ 的状态可以由某个随机变量 $x(t,\boldsymbol{\xi},\omega)$ 来表达的情形. 现设

$$m(t,\boldsymbol{\xi})=E(x(t,\boldsymbol{\xi},\omega)),$$
$$v(t,\boldsymbol{\xi};s,\boldsymbol{\eta})=E\{(x(t,\boldsymbol{\xi},\omega)-m(t,\boldsymbol{\xi}))\overline{(x(s,\boldsymbol{\xi},\omega)-m(s,\boldsymbol{\xi}))}\}.$$

若这对时间、空间的平行移动都不变, 则 $x(t,\boldsymbol{\xi},\omega)$ 是平衡的. 此时, $m(t,\boldsymbol{\xi})=$ 常数 (以后令为 0), 而 $v(t,\boldsymbol{\xi};s,\boldsymbol{\eta})$ 就成为 $t-s,\boldsymbol{\xi}-\boldsymbol{\eta}=(\xi_1-\eta_1,\xi_2-\eta_2,\boldsymbol{\xi}_3-\boldsymbol{\eta}_3)$ 的函数 $v(t-s,\boldsymbol{\xi}-\boldsymbol{\eta})$. 关于 v 亦可得对应于 Khinchin 分解的谱

分解 ①.

$$v(t,\boldsymbol{\xi})=\int_{\mathbb{R}^4}\mathrm{e}^{-i2\pi(\lambda t+(\delta,\boldsymbol{\xi}))}\mathrm{d}F(\lambda,\delta),$$
$$(\sigma,\boldsymbol{\xi})=\sigma_1\xi_1+\delta_2\xi_2+\sigma_3\xi_3.$$

而 $x(t,\boldsymbol{\xi},\omega)$ 的分解 (Kolmogoroff 分解) 也与平稳过程的场合一样.

又除对地点的齐次性以外, 还可假定各向同性的性质. 这时, $v(t,\boldsymbol{\xi})$ 可以写成 $v(t,|\boldsymbol{\xi}|)$ 的形状, $|\boldsymbol{\xi}|$ 是向量 $\boldsymbol{\xi}$ 的长度, 此时测度 $\mathrm{d}F(\lambda,\sigma)$ 对地点也具有各向同性的性质, 即 $\mathrm{d}F(\lambda,r)=\mathrm{d}G(\lambda,r)\cdot\mathrm{d}\theta$($r$ 表示 σ 的长度, 而 θ 就表示 σ 切单位球的点). 也就是

$$v(t,|\xi|)=\int_{\mathbb{R}^2}\mathrm{e}^{-i2\pi\lambda t}K(r\cdot|\xi|)\mathrm{d}G(\lambda,r). \tag{31.14}$$

此处 K 可用 Bessel 函数表为

$$K(p)=2\pi\frac{J_{1/2}(2\pi p)}{p^{1/2}}. \tag{31.15}$$

各向同性湍流(isotropic turbulence) 前面说的是对时点 t 和空间的点 $\boldsymbol{\xi}$, 对应着一个偶然的纯量 $x(t,\boldsymbol{\xi},\omega)$, 但在这里所考虑的是以向量 $\boldsymbol{u}(t,\boldsymbol{\xi},\omega)$ 代替纯量. 譬如, 对任意的时点 t, 令湍流在空间的点 $\boldsymbol{\xi}$ 上的速度为 $\boldsymbol{u}(t,\boldsymbol{\xi},\omega)$ 就是这种情形. 为了弄清楚实质上的难点, 而把时点 t 固定起来, 并记为 $\boldsymbol{u}(\boldsymbol{\xi},\omega)$, 再假定 $E\boldsymbol{u}(\boldsymbol{\xi},\omega)=0$. 于是 $v(\boldsymbol{\xi},\boldsymbol{\eta})$ 就是

$$v(\boldsymbol{\xi},\boldsymbol{\eta})=E[\boldsymbol{u}(\boldsymbol{\xi},\omega)\otimes\boldsymbol{u}(\boldsymbol{\eta},\omega)],$$

而这就是在点 $\boldsymbol{\xi},\boldsymbol{\eta}$ 上的切平面 $T_{\boldsymbol{\xi}}$ 和 $T_{\boldsymbol{\eta}}$ 的张量积 $T_{\boldsymbol{\xi}}\otimes T_{\boldsymbol{\eta}}$ 的元素. 若这对于空间的全等变换为不变, 则 $\boldsymbol{u}(\boldsymbol{\xi},\omega)$ 叫做各向同性湍流. 这个不变性的意义, 讲得清楚些就是这样的: 令 g 为任意的全等变换, g 将点 $\boldsymbol{\xi}$ 映照成点 $g\cdot\boldsymbol{\xi}$, 而同时由此引出 $\boldsymbol{\xi}$ 的切平面 $T_{\boldsymbol{\xi}}$ 到 $g\cdot\boldsymbol{\xi}$ 的切平面 $T_{g\cdot\boldsymbol{\xi}}$ 上的全等变换 $\dot{g}$, 还引出 $T_{\boldsymbol{\xi}}\otimes T_{\boldsymbol{\eta}}$ 到 $T_{g\cdot\boldsymbol{\xi}}\otimes T_{g\cdot\boldsymbol{\eta}}$ 上的全等变换. 后者也以 $\dot{g}$ 来表示. 由 g 得出的不变性就是

$$E[\dot{g}\boldsymbol{u}(\xi,\omega)\otimes\dot{g}\boldsymbol{u}(\boldsymbol{\eta},\omega)]=E[\boldsymbol{u}(g\boldsymbol{\xi},\omega)\otimes\boldsymbol{u}(g\cdot\eta,\omega)]. \tag{31.16}$$

① 这是多元的 Khinchin-Bochner 定理, 它的证明可参见 S. Bochner, Harmonic Analysis and the Theory of Probability, 1955, 58 页. —— 校者注

也就是

$$gv(\boldsymbol{\xi},\boldsymbol{\eta})=v(g\boldsymbol{\xi},g\boldsymbol{\eta}).$$

现在, 若在 $\boldsymbol{\xi}$ 的空间里定一正交系 (e_1,e_2,e_3), 又在 $T_{\boldsymbol{\xi}}$ 里也取一与这平行的正交系 $(e_1(\boldsymbol{\xi}),e_2(\boldsymbol{\xi}),e_3(\boldsymbol{\xi}))$, 然后以此来决定坐标, 就得

$$\left.\begin{aligned}&\boldsymbol{\xi}'=g\boldsymbol{\xi}\Leftrightarrow\xi_i'=\sum_j g_{ij}\xi_j+h_i,(g_{ij})=\text{正交矩阵}\\&\boldsymbol{u}'=\dot{g}\cdot\boldsymbol{u}\Leftrightarrow u_i'=\sum_j g_{ij}u_j,\end{aligned}\right\}\tag{31.17}$$

因而 (31.16) 就可写成

$$E\left(\sum_k g_{ik}u_k(\boldsymbol{\xi})\sum_l g_{jl}u_l(\boldsymbol{\eta})\right)=E(u_i(g\boldsymbol{\xi})\cdot u_j(g\boldsymbol{\eta})),$$

也就是

$$\sum_{kl}g_{ik}g_{jl}v_{kl}(\boldsymbol{\xi},\boldsymbol{\eta})=v_{ij}(g\boldsymbol{\xi},g\boldsymbol{\eta}).\tag{31.18}$$

特别地, 若设 $(g_{ij})=$ 单位矩阵, 则

$$v_{ij}(\boldsymbol{\xi},\boldsymbol{\eta})=v_{ij}(\boldsymbol{\xi}+\boldsymbol{h},\boldsymbol{\eta}+\boldsymbol{h}),\quad \boldsymbol{h}=(h_1,h_2,h_3).$$

故 v_{ij} 是 $\boldsymbol{\xi}$-$\boldsymbol{\eta}$ 的函数. 若记这函数为 $v_{ij}(\boldsymbol{\xi}\text{-}\boldsymbol{\eta})$, 则由 (31.18) 得

$$\sum_{kl}g_{ik}g_{jl}v_{kl}(\boldsymbol{\xi})=v_{ij}(g\boldsymbol{\xi}).$$

现在, 若设 $v(\boldsymbol{\xi};a,b)=\sum\limits_{ij}a_ib_jv_{ij}(\boldsymbol{\xi})$, 则上式等价于

$$v(g\boldsymbol{\xi};ga,gb)=v(\boldsymbol{\xi};a,b),\quad g\text{是正交矩阵}.\tag{31.19}$$

因由定义, $v(\boldsymbol{\xi};a,a)$ 是 $\boldsymbol{\xi}$ 的正定函数, 所以

$$v(\boldsymbol{\xi};a,a)=\int \mathrm{e}^{-i2\pi(\lambda,\boldsymbol{\xi})}m(\mathrm{d}\lambda;a),\quad m(\mathrm{d}\lambda;a)\geqslant 0.$$

利用 $v(\boldsymbol{\xi};a,b)$ 是 a,b 的双二次形式这一事实, 就得

$$\begin{aligned}v(\boldsymbol{\xi};a,a)&=\int \mathrm{e}^{-i2\pi(\lambda,\boldsymbol{\xi})}m(\mathrm{d}\lambda;a,b),\\m(\mathrm{d}\lambda;a,a)&=m(\mathrm{d}\lambda;a)\geqslant 0.\end{aligned}\tag{31.20}$$

$m(\mathrm{d}\lambda;a,b)$ 就是 a,b 的正定的双二次形式. 又由 v 的不变性 (31.19), 便知 m 也具有不变性:

$$m(g\cdot\mathrm{d}\lambda;ga,gb)=m(\mathrm{d}\lambda;a,b).$$

由此就得出表达式

$$\begin{aligned}m(\mathrm{d}\lambda;a,b)=&\sum a_ib_j[\theta_i\theta_j\mathrm{d}\theta m_1(\mathrm{d}r)\\&+(\delta_{ij}-\theta_i\theta_j)\mathrm{d}\theta m_2(\mathrm{d}r)+m_0(\mathrm{d}\lambda)].\end{aligned}\tag{31.21}$$

此处 $r=|\lambda|,\theta_i=\lambda_i/|\lambda|,\mathrm{d}\theta$ 是单位球上的球面元素,m_1,m_2 是 $(0,\infty)$ 上的有界测度, m_0 是只分布于 λ 空间的原点的测度. 反之, 若由形如 (31.21) 的 m 和 (31.20) 来定义 v, 便知这是对应于各向同性湍流的 v.

第 4 章　Markoff 过程

§32　条 件 概 率

考虑概率空间 $\Omega(\boldsymbol{B}, P)$. 设 $\boldsymbol{B}_1$ 为 $\boldsymbol{B}$ 的子 Borel 集合体. 对 $\boldsymbol{B}_1$ 可测的集或函数自然对 $\boldsymbol{B}$ 也是可测的, 但反之未必成立. 次设 A 为事件, 即对 $\boldsymbol{B}$ 可测的集. 按照 Doob 的说法, 可将 A 关于 $\boldsymbol{B}_1$ 的**条件概率** (conditional probability)$P(A/\boldsymbol{B}_1)$ 定义为满足下列条件的 ω 的实函数 $P(A/\boldsymbol{B}_1)(\omega)$:

(C.1) $P(A/\boldsymbol{B}_1)(\omega)$ 对 $\boldsymbol{B}_1$ 可测 (故当然是 $\Omega(\boldsymbol{B}, P)$ 上的随机变量).

(C.2) 若 B 为 $\boldsymbol{B}_1$ 的可测集, 即 $B \in \boldsymbol{B}_1$, 则

$$P(A \cdot B) = \int_B P(A/\boldsymbol{B}_1)(\omega)P(\mathrm{d}\omega). \tag{32.1}$$

利用 Radon-Nikodym 定理可以证明上述的 $P(A/\boldsymbol{B}_1)(\omega)$ 是唯一存在的 (除零测集外). 因若将 $P(A \cdot B)$ 看成为关于 B 的集涵数, 则它是 $\Omega(\boldsymbol{B}_1)$ 上的有限测度, 并由 $P(A \cdot B) \leqslant P(B)$, 知它关于 P(精确些, 限制在 $\boldsymbol{B}_1$ 上的 P) 为绝对连续, 故满足 (32.1), 对 $\boldsymbol{B}_1$ 可测的函数除在零测集上外是唯一的.

因易使人怀疑上述定义与通常所定义的条件概率究竟有何关系, 故就此略加说明. 设将 Ω 分成为有限或可数个互不相交的可测集之和, 即

$$\Omega = B_1 + B_2 + \cdots \tag{32.2}$$

将 $B_1, B_2, \cdots$ 中若干个 (0 个, 有限个, 无限个) 集的和的集的全体表为 $\boldsymbol{B}_1$, 则 $\boldsymbol{B}_1$ 显然是一 Borel 集合体. 此时对 $\boldsymbol{B}_1$ 可测的函数 $P(A/\boldsymbol{B}_1)(\omega)$ 分别在 $B_1, B_2, \cdots$ 上各取常数值 $a_1, a_2, \cdots$, 故在 (32.1) 中, 若以 B_i 代 B, 则得

$$P(A \cdot B_i) = a_i P(B_i),$$

亦即

$$a_i = P(A \cdot B_i)/P(B_i).$$

故得下式

$$P(A/\boldsymbol{B}_1)(\omega) = P(A\cdot B_i)/P(B_i), \quad \omega\in B_i. \tag{32.3}$$

通常 $P(AB)/P(B)$ 叫做在条件 B 下的 A 的概率, 并记为 $P(A/B)$. 利用这一记号, (32.3) 变成

$$P(A/\boldsymbol{B}_1)(\omega) = P(A/B_i), \quad \omega\in B_i. \tag{32.3'}$$

由此就显示了 Doob 定义与通常的定义的关系.

次设 $x(\omega)$ 为一随机向量, 其值域为 $\mathbb{R}^\Lambda(\boldsymbol{B}^\Lambda)$. 设 $\boldsymbol{B}_1$ 为 $\{x^{-1}(E)/E\in\boldsymbol{B}^\Lambda\}$. 当 $x(\omega)$ 的坐标为 $x_\lambda(\omega)$ 时, 可以认为 $\boldsymbol{B}_1$ 是由形如

$$x_\lambda(\omega)<c$$

的集所产生的 Borel 集合体. 由于 $\mathbb{R}^\Lambda(\boldsymbol{B}^\Lambda)$ 上 $(\boldsymbol{B}^\Lambda)$ 可测函数 φ 可写成 $\varphi(x(\omega))$ 的形式, 故由 $P(A/\boldsymbol{B}_1)(\omega)$ 对 $\boldsymbol{B}_1$ 可测的条件, 可将 (32.1) 写成

$$P(A\cdot x^{-1}(E)) = \int_{x^{-1}(E)}\varphi(x(\omega))P(\mathrm{d}\omega).$$

若令 x 的分布为 P_x, 则因 $P_x = P\cdot x^{-1}$, 故得

$$P(A\cdot x^{-1}(E)) = \int_E \varphi(\xi)P_x(\mathrm{d}\xi).$$

上式左边可看成 $E\in\boldsymbol{B}^\Lambda$ 的函数, 它是一测度, 且因 $P(A\cdot x^{-1}(E))\leqslant P(x^{-1}(E)) = P_x(E)$, 故它是关于 P_x 绝对连续的测度. 因此, 以上条件唯一地决定了 $\varphi(\xi)$, 此即 Kolmogoroff 所定义的条件概率 $P(A/x=\xi)$. 前面 Doob 所定义的 $P(A/\boldsymbol{B}_1)(\omega)$ 就是将 $x(\omega)$ 代入 $\varphi(\xi)$ 中的 ξ 而得的随机变量, 以后简写为 $P(A/x)$.

例 设 x, y 为两个实随机变量, 并设其联合分布有密度 $f(\xi,\eta)$, 则 Kolmogoroff 条件概率为

$$P(y\in E/x=\xi) = \int_E f(\xi,\eta)\mathrm{d}\eta\Big/\int_{\mathbb{R}^1} f(\xi,\eta)\mathrm{d}\eta.$$

由此可得 Doob 条件概率为

$$P(y\in E/x) = \int_E f(x,\eta)\mathrm{d}\eta\Big/\int_{\mathbb{R}^1} f(x,\eta)\mathrm{d}\eta.$$

§33 条件数学期望

条件期望完全可以和条件概率并行地定义. 设 $y(\omega)$ 为实 (或复) 随机变量, 并设

$$E|y| < \infty. \tag{33.1}$$

与上节一样, 设 $\boldsymbol{B}_1$ 为 $\boldsymbol{B}$ 的子 Borel 体. 对任意可测集 B, 定义

$$E(y;B) = \int_B y(\omega)P(\mathrm{d}\omega). \tag{33.2}$$

y 对于 $\boldsymbol{B}_1$ 的条件期望 $E(y/\boldsymbol{B}_1)$ 定义为满足以下两个条件的 ω 的函数 $E(y/\boldsymbol{B}_1)(\omega)$.

(E.1) $E(y/\boldsymbol{B}_1)(\omega)$ 对 $\boldsymbol{B}_1$ 可测.

(E.2) 对任意 $B \in \boldsymbol{B}_1, E(y;B) = E(E(y/\boldsymbol{B}_1);B)$.

这样的 $E(y/\boldsymbol{B}_1)(\omega)$ 必存在, 且除 P 零测集外是唯一的, 这可与条件概率的情形一样, 利用 Radon-Nikodym 定理来证明.

若设 A 的示性函数为 $\chi_A(\omega)$, 由于 $E(\chi_A/\boldsymbol{B}_1)$ 与 $P(A/\boldsymbol{B}_1)$ 一致, 故条件概率可看成条件期望的特殊情形. 下面叙述有关条件期望的性质, 由此亦易推出条件概率的性质.

(i) 若 z 对 $\boldsymbol{B}_1$ 可测, 且 $E|y| < \infty, E|yz| < \infty$, 则

$$E(zy/\boldsymbol{B}_1) = zE(y/\boldsymbol{B}_1). \tag{33.3}$$

因此, 如 z 为 $M \in \boldsymbol{B}_1$ 的示性函数, 则对任意 $B \in \boldsymbol{B}_1$, 有

$$\begin{aligned} E(zy;B) = E(y;M\cdot B) &= \int_{M\cdot B} E(y/\boldsymbol{B}_1)(\omega)P(\mathrm{d}\omega) \\ &= \int_B z(\omega)E(y/\boldsymbol{B}_1)(\omega)P(\mathrm{d}\omega). \end{aligned}$$

并由 $z(\omega)E(y/\boldsymbol{B}_1)(\omega)$ 对 $\boldsymbol{B}_1$ 可测, 得

$$E(zy/\boldsymbol{B}_1)(\omega) = z(\omega)E(y/\boldsymbol{B}_1)(\omega).$$

再注意 (33.3) 的两边关于 z 是线性的, 由此即推出对任意的 z, (33.3) 均成立.

(ii) 若 $\boldsymbol{B}_1 \subset \boldsymbol{B}_2$, 则

$$E(y/\boldsymbol{B}_1) = E(E(y/\boldsymbol{B}_2)/\boldsymbol{B}_1), \tag{33.4}$$

因若 $B \in \boldsymbol{B}_1$, 则显然 $B \in \boldsymbol{B}_2$, 故

$$\begin{aligned} E(y; B) &= E\{E(y/\boldsymbol{B}_2); B\} \\ &= E\{E(E(y/\boldsymbol{B}_2)/\boldsymbol{B}_1); B\}. \end{aligned}$$

故关系式 (33.4) 成立.

(iii) 如 $y = c_1 y_1 + c_2 y_2$ (c_1, c_2 为常数), $E|y_i| < \infty, i = 1, 2$, 则

$$E(y/\boldsymbol{B}_1) = c_1 E(y_1/\boldsymbol{B}_1) + c_2 E(y_2/\boldsymbol{B}_1). \tag{33.5}$$

(iv) 如 $y_n \to y, |y_n| \leqslant z, E|z| < \infty$, 则

$$E(y_\eta/\boldsymbol{B}_1) \to E(y/\boldsymbol{B}_1). \tag{33.6}$$

注意 1 以上关系除一 P 零测集外皆成立.

注意 2 若 $\boldsymbol{x}$ 为任意的随机向量, 则与 $P(A/x)$ 同样可以定义 $E(y/\boldsymbol{x})$.

§34 鞅①

鞅 (Martingale) 在概率论中是一个非常重要的概念, 但在本节中只叙述与本章有关的内容.

设 $x_t, t \in T$(T 为实数集), 为一随机过程, 且 $E(|x_t|) < \infty$, 设对每一 $t \in T$ 有一 Borel 集合体 $\boldsymbol{B}_t (\subseteq \boldsymbol{B})$ 与之对应, 且设

(i) 如 $s < t$, 则 $\boldsymbol{B}_s \subset \boldsymbol{B}_t$;

(ii) 对所有 $t \in T, x_t$ 对 $\boldsymbol{B}_t$ 可测;

(iii) 如 $s < t$, 则 $E\{x_t/\boldsymbol{B}_s\} = xs$ 以概率为 1 地成立.

此时称 $\{x_t\}$ 关于 $\{\boldsymbol{B}_t\}$ 为鞅.

如条件 (iii) 换为

(iii′) 若 $s < t$, 则 $E\{x_t/\boldsymbol{B}_s\} \leqslant x_s$ 以概率为 1 地成立,

就称 $\{x_t\}$ 关于 $\{\boldsymbol{B}_t\}$ 为半鞅 (Semi-Martingale). 鞅是半鞅的特殊情形.

试述半鞅的性质. 首先, 下式成立:

① 此节可参看 J. L. Doob, Stochastic Processes, 1953, 第 7 章 §11. —— 校者注

$$\text{如 } s<t, \text{ 则 } E(x_s)\leqslant E(x_t).$$

此式易由 (iii′) 导出. $E(x_t)$ 既然对 t 单调不减, 故至多除在可数多个点外, 均为连续.

定理 1 设 T 为区间, $x_t, t\in T$, 关于 $B_t(t\in T)$ 为半鞅, 且 $E(x_t)$ 连续. 若 x_t 为可分, 则几乎一切样本函数 ω 仅有第一类不连续点, 且对每一 t, 有

$$P(x_{t-0}=x_t=x_{t+0})=1.$$

(证略)

§35 转移概率

设 R 为具有第二可数性的紧致的 Hausdorff 空间, 因而 R 的拓扑可由距离决定. 令 $\boldsymbol{B}_R$ 为含 R 的一切开集的最小 Borel 体.

函数 $P(t,x,E)$, 其中 t 为时间, $t\geqslant 0, x\in R, E\in \boldsymbol{B}_R$, 如满足下列条件, 则称为**转移概率** (transition probability):

(T.1) 当 t,x 固定时, $P(t,x,E)$ 关于 E 是概率分布.

(T.2) (对 x 的连续性) 对固定的 t, 当 $x\to x_0$ 时, 测度 $P(t,x,E)$ 弱收敛于 $P(t,x_0,E)$; 即对任意的连续函数 f,

$$\int_R f(y)P(t,x,\mathrm{d}y)\to\int_R f(y)P(t,x_0,\mathrm{d}y),\quad x\to x_0.$$

(T.3) (对 t 的连续性) 对固定的 x, 当 $t\to 0$ 时, 测度 $P(t,x,E)$ 弱收敛于 $\delta(x,E)(=1, x\in E; =0, x\in E^c)$; 即对于任意的连续函数 f,

$$\int_R f(y)P(t,x,\mathrm{d}y)\to\int_R f(y)\delta(x,\mathrm{d}y)\equiv f(x),\quad t\to 0.$$

注意, 由 (T.2) 可以证明, 当固定 t 和 E 时, $P(t,x,E)$ 是 x 的 $(\boldsymbol{B}_R)$ 可测函数. 为此, 只要证明对于 $(\boldsymbol{B}_R)$ 任意有界可测函数 f,

$$\varphi(x)=\int_R f(y)P(t,x,\mathrm{d}y)$$

关于 x 是 $(\boldsymbol{B}_R)$ 可测函数即可. 实际上, 记具有这种性质的 f 的全体为 $\mathfrak{F}$, 由 (T.2) 知 $\mathfrak{F}$ 包含一切连续函数. 又由可测性的定义知 $\mathfrak{F}$ 对极限运算不变, 故任意有界 $(\boldsymbol{B}_R)$ 可测函数皆属于 $\mathfrak{F}$.

注意上述各点后, 再加上下面的条件:

(T.4) **(Chapman-Kolmogoroff 方程)**

$$P(t+s,x,E)=\int_R P(t,x,\mathrm{d}y)P(s,y,E).$$

由上面的叙述可知右边的积分有意义.

$P(t,x,E)$ 的直观意义是: 在最初处于 x 状态的条件下, 经过 t 时后转移到 E 状态的概率. 因此 Chapman-Kolmogoroff 方程的意义是: 为了经时间 $t+s$ 后自 x 转移到 E, 先经 t 时后转入 R 的某点 y, 然后再经时间 s 后转入 E. 这个方程是对这种转移必然要求的条件.

例 1　设 R 为有限集. 显然 R 对离散拓扑 (discrete topology) 而言满足上述条件. 为了定义 $P(t,x,E)$, 只要对点集 $\{y\}$ 定义了 E 即可, 记为 $P(t,x,y)$. 由 (T.1) 得

$$P(t,x,y)\geqslant 0,\quad \sum_y P(t,x,y)=1. \tag{35.1}$$

(T.2) 显然成立.(T.3) 变成

$$P(t,x,y)\to\delta(x,y)=\begin{cases}1, & x=y,\\ 0, & x\neq y.\end{cases} \tag{35.2}$$

当 t 固定时, $(P(t,x,y),x,y\in R)$ 可视为有限阶矩阵, 记为 $\boldsymbol{P}_t$, 则 (35.1) 表示 $\boldsymbol{P}_t$ 的元素非负, 且其各行元素的和为 1. 这种矩阵称为**随机矩阵** (stochastic matrix). (35.2) 表示着

$$\boldsymbol{P}_t\to\boldsymbol{I}(\text{单位矩阵}),t\to 0. \tag{35.2$'$}$$

由矩阵乘法, Chapman-Kolmogoroff 方程变为

$$\boldsymbol{P}_{t+s}=\boldsymbol{P}_t\cdot\boldsymbol{P}_s. \tag{35.3}$$

例 2　设 R 为实数集 $\mathbb{R}^1$ 加上 ∞ 的紧化空间. 对 $x\in\mathbb{R}^1,E\in\mathbb{R}^1$, 设

$$P(t,x,E)=\int_E N_t(y-x)\mathrm{d}y,\quad N_t(x)=\frac{1}{\sqrt{2\pi t}}\mathrm{e}^{-\frac{x^2}{2t}}.$$

再设

$$P(t,\infty,E)=\delta(\infty,E).$$

(T.1) 显然成立. 试研究对于 x 的连续性. $x_0\neq\infty$ 时, 对于连续函数 f,

$$\int f(y)N_t(y-x)\mathrm{d}y=\int f(y+x)N_t(y)\mathrm{d}y$$
$$\to\int f(y+x_0)N_t(y)\mathrm{d}y=\int f(y)N_t(y-x_0)\mathrm{d}y,$$

(因 f 在紧空间 $R=\mathbb{R}^1\cup\{\infty\}$ 上连续, 故有界). 当 $x_0=\infty$ 时, 对所有的 y 可得 $f(y+x)\to f(\infty)$ $(x\to\infty)$, 故与上式同样可得

$$\int f(y)N_t(y-x)\mathrm{d}y\to f(\infty)=\int f(y)\delta(\infty,\mathrm{d}y).$$

对于 t 的连续性可导出

$$\int f(y)N_t(y-x)\mathrm{d}y=\int f(x+\sqrt{t}y)N_1(y)\mathrm{d}y\underset{t\to 0}{\longrightarrow}f(x).$$

Chapman-Kolmogoroff 方程在 $\mathbb{R}^1$ 中可由正态分布的性质 $N_{t+s}=N_t*N_s$ 导出, 在 ∞ 处则显然成立.

例 3 设 $R=[0,\infty]$, 并设 $P(t,x,E)$ 由下式定义:

$$P(t,x,E)=\int_E[N_t(y-x)+N_t(y+x)]\mathrm{d}y, x\in[0,\infty), E\subset[0,\infty),$$
$$P(t,\infty,E)=\delta(\infty,E).$$

容易验证, $P(t,x,E)$ 满足上述四个条件 (T.1)~(T.4).

注意 利用 (T.4) 可自 (T.3) 推得 (参看 §37):

(T.3′) 对固定的 x, 当 $t\to t_0$ 时, $P(t,x,E)$ 弱收敛于 $P(t_0,x,E)$.

由此可见为何 (T.3) 可表达对 t 的连续性.

§36 伴随转移概率的半群与对偶半群①

利用前节的记号, 记在 R 上的连续函数的全体为 C. 对通常的线性运算来说, C 是线性空间, 赋与范数 $\|f\|=\max|f(z)|, x\in R$, 则它可看作可分的 Banach 空间. 今定义 $T_t(t>0)$ 如下:

$$T_tf(x)=\int_R f(y)P(t,x,\mathrm{d}y), \tag{36.1}$$

① 关于本节和下节, 可参看关肇直编的《泛函分析讲义》, 高等教育出版社, 1958, 第 4 章. ——校者注

则由 (T.2) 可知, 如 $f \in \boldsymbol{C}$, 则 $T_t f \in \boldsymbol{C}$, 且 T_t 可看做 $\boldsymbol{C}$ 中的线性算子. 利用 (T.1) 可得

$$T_t \geqslant 0, \text{ 即如 } f \geqslant 0, \text{ 则 } T_t f \geqslant 0, \tag{36.2}$$

$$T_t 1 = 1, \tag{36.3}$$

因而

$$\| T_t \| \equiv \sup\{\| T_t f \|, \| f \| \leqslant 1\} = 1. \tag{36.4}$$

由 (T.3) 得

$$T_t f(x) \to f(x), \quad x \in R, t \to 0. \tag{36.5}$$

因 $\boldsymbol{C}$ 的共轭空间 $\boldsymbol{C}^*$ 为 R 上的 (有符号的) 测度空间, 故 (36.5) 式变成

$$T_t \to I(\text{弱}) \quad (I \text{ 为恒等算子}). \tag{36.5'}$$

亦即对任意的 $f \in \boldsymbol{C}, \mu \in \boldsymbol{C}^*$, 上式等价于

$$(T_t f, \mu) \to (f, \mu).$$

由 (T.4) 可得半群的性质

$$T_t T_s = T_{t+s}. \tag{36.6}$$

一般地, 对于可分 Banach 空间 E 上的有界线性算子族 $T_t, t > 0$, 如有性质

$$\| T_t \| \leqslant 1, \quad T_t \to I(\text{弱}), \quad T_t T_s = T_{t+s}, \tag{36.7}$$

则称此族为 E 上算子的半群, 或简称为 E 上的**半群** (semi-group).

上述 $\boldsymbol{C}$ 上的算子族 $T_t, t > 0$, 亦满足 (36.7), 故此族是 $\boldsymbol{C}$ 上的半群, 称为**伴随 $\boldsymbol{P(t, x, E)}$ 的半群.**

其次, 对 $\mu \in \boldsymbol{C}^*$, 定义

$$T_t^* \mu(E) = \int_R P(t, x, E) \mu(\mathrm{d}x), \tag{36.8}$$

于是由 $0 \leqslant P(t, x, E) \leqslant 1$, 得

$$\| T_t^* \mu \| \leqslant \| \mu \|, \quad \text{即} \quad \| T_t^* \| \leqslant 1, \tag{36.9}$$

此处 $\| \mu \|$ 是 μ 的全变差 (total variation). 注意到

$$(T_t f, \mu) = (f, T_t^* \mu) = \int_R \int_R f(y) P(t, x, \mathrm{d}y) \mu(\mathrm{d}x),$$

可见 T_t^* 是 T_t 的共轭算子, 故记号 T_t^* 是合理的. 由 Chapman-Kolmogoroff 方程得 $T_t^* T_s^* = T_{t+s}^*$. 又

$$(f, T_t^* \mu) = (T_t f, \mu) \to (f, \mu), f \in \boldsymbol{C}, \mu \in \boldsymbol{C}^*.$$

因 $\boldsymbol{C}$ 不等于 $(\boldsymbol{C}^*)^*$, 故不能导出 $T_t^* \to I^*$(弱)(I^* 为 $\boldsymbol{C}^*$ 上的恒等算子), 但因它是与此相似的条件, 故称为 $T_t^* \to I^*$(泛弱).

综合上述得

$$\| T_t^* \| \leqslant 1, \quad T_t^* \to I^*(\text{泛弱}), \quad T_t^* T_s^* = T_{t+s}^*. \tag{36.10}$$

因 $\boldsymbol{C}^*$ 不是可分的, 而且上述第二条件与 (36.7) 的第二条件稍有差异, 故不能说 $T_t^*(t > 0)$ 是 $\boldsymbol{C}^*$ 上的半群, 但因比较近似, 故称为伴随 $P(t, x, E)$ 的**对偶半群** (dual semi-group).

§37 Hille-Yosida 理论 (1)

设 E 为可分 Banach 空间, $T_t(t > 0)$ 为其上的半群. 在前节 (36.7) 中只要求 $T_t \to I$(弱), 但利用其他性质后, 可得

$$\| T_t f - f \| \to 0, \quad \text{即 } T_t \to I(\text{强})^{①}. \tag{37.1}$$

要证明这个性质, 需要用到 Dunford 定理 (Ann. of Math., 33, 1932, pp.567~573), 故证明从略. 由 (37.1) 还可得到

$$\| T_t f - T_s f \| \to 0 \quad (t \to s). \tag{37.2}$$

为此, 设 $u = \min(t, s)$, $v = |t - s|$, 则得

$$\| T_t f - T_s f \| = \| T_u (T_v f - f) \| \leqslant \| T_v f - f \| \to 0.$$

由此可知 (37.2) 的收敛对 s 是一致的, 故 $T_t f$ 对 t 为一致连续.

① 参见关肇直编的《泛函分析讲义》, 高等教育出版社, 1958, 482 页. —— 校者注

Hille-Yosida 理论是关于半群的生成算子的理论. 由于 $T_t(t>0)$ 有群的性质 $T_tT_s=T_{t+s}$, 因此有时虽未必尽知一切的 $T_t, t>0$, 但如知道 $T_t, 0<t<\delta$, 此处 δ 为任意小的正数, 则对任意 t, 都可求出 T_t. $T_t(0<t<\delta)$ 称为半群 $T_t(t>0)$ 的**群芽** (germ). 为了给出 $\delta\to 0$ 时的极限性质, 定义算子 A 使

$$Af=\lim_{t\downarrow 0}\frac{T_tf-f}{t}\quad (\text{极限对范数而取}), \tag{37.3}$$

A 称为**生成算子** (generator). 使上述极限存在的 f 全体形成 A 的定义域 $\mathfrak{D}(A)$.

显然 $\mathfrak{D}(A)$ 为 E 的线性子空间, A 为线性算子. 通常 $\mathfrak{D}(A)$ 与 E 并不一致, 且 A 非有界. 但为了易于说明, 先对 A 为有界且 $\mathfrak{D}(A)=E$ 的情形作出形式的计算. 设对于算子的范数而言, 有

$$A=\lim_{t\downarrow 0}\frac{T_t-I}{t}. \tag{37.4}$$

由

$$\frac{T_{t+\delta-T_t}}{\delta}=\frac{T_\delta-I}{\delta}T_t$$

当 $\delta\to 0$ 时得

$$\frac{\mathrm{d}T_t}{\mathrm{d}t}=AT_t.$$

与普通数的情形一样, 解上式得

$$T_t=\mathrm{e}^{tA}\Big(=\sum_{n=0}^{\infty}\frac{(tA)^n}{n!}\Big).$$

考虑 T_t 的 Laplace 变换 R_λ

$$R_\lambda=\int_0^\infty \mathrm{e}^{-\lambda t}T_t\mathrm{d}t\Big(=\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{\infty}\mathrm{e}^{-\lambda\frac{k}{n}}T_{\frac{k}{n}}\Big).$$

由形式的计算得

$$R_\lambda=\int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{e}^{tA}\mathrm{d}t=\int_0^\infty \mathrm{e}^{-t(\lambda I-A)}\mathrm{d}t=(\lambda I-A)^{-1}.$$

因 $\lambda If \equiv \lambda f$, 故 λI 可简写为 λ,

$$R_\lambda = (\lambda - A)^{-1}.$$

即 $u = R_\lambda v$ 满足

$$(\lambda - A)u = v,$$

根据上述定义, R_λ 称为 T_t 的预解算子 (resolvent).

留意上述各点, 今试就一般情形加以严格论证.

预解算子 R_λ 由下式定义:

$$(R_\lambda f, \mu) = \int_0^\infty \mathrm{e}^{-\lambda t}(T_t f, \mu)\mathrm{d}t, \quad f \in E, \mu \in E^*. \tag{37.5}$$

亦可以定义为

$$R_\lambda f = \int_0^\infty \mathrm{e}^{-\lambda t} T_t f \mathrm{d}t \Big(= \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda \frac{k}{n}} T_{\frac{k}{n}} f \Big). \tag{37.5'}$$

如上所述, 因 $T_t f$ 对 t 连续而且有界, 故以上定义是可能的.

$$\| R_\lambda f \| \leqslant \int_0^\infty \mathrm{e}^{-\lambda t} \| T_t f \| \,\mathrm{d}t \leqslant \| f \| \int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{d}t = \frac{\| f \|}{\lambda}.$$

故得

$$\| R_\lambda \| \leqslant \frac{1}{\lambda}, \tag{37.6}$$

即 R_λ 为有界线性算子.

其次, 再证

$$R_\lambda - R_\mu = -(\lambda - \mu) R_\lambda R_\mu, \tag{37.7}$$

$$\| \lambda R_\lambda f - f \| \to 0, \quad \lambda \to \infty. \tag{37.8}$$

对于 $f \in E, \sigma \in E^*$,

$$\begin{aligned}(R_\lambda R_\mu f, \sigma) &= \int_0^\infty \mathrm{e}^{-\lambda t}(T_t R_\mu f, \sigma)\mathrm{d}t \\ &= \int_0^\infty \mathrm{e}^{-\lambda t}(R_\mu f, T_t^* \sigma)\mathrm{d}t \\ &= \int_0^\infty \mathrm{e}^{-\lambda t} \int_0^\infty \mathrm{e}^{-\mu s}(T_s f, T_t^* \sigma)\mathrm{d}s\mathrm{d}t\end{aligned}$$

$$\begin{aligned}
&= \int_0^\infty e^{-\lambda t} \int_0^\infty e^{-\mu s}(T_{s+t}f, \sigma)\mathrm{d}s\mathrm{d}t \\
&= \int_0^\infty e^{-\lambda t+\mu t} \int_t^\infty e^{-\mu s}(T_s f, \sigma)\mathrm{d}s\mathrm{d}t \\
&= \int_0^\infty e^{-\mu s}(T_s f, \sigma)\mathrm{d}s \int_0^s e^{-(\lambda-\mu)t}\mathrm{d}t \\
&= \int_0^\infty e^{-\mu s}(T_s f, \sigma)\frac{e^{-(\lambda-\mu)s}-1}{-(\lambda-\mu)}\mathrm{d}t \\
&= \frac{1}{-(\lambda-\mu)}\int_0^\infty (e^{-\lambda s}-e^{-\mu s})(T_s f, \sigma)\mathrm{d}t \\
&= \frac{1}{-(\lambda-\mu)}[(R_\lambda - R_\mu)f, \sigma].
\end{aligned}$$

$$\| \lambda R_\lambda f - f \| \leqslant \int_0^\infty \lambda e^{-\lambda t} \| T_t f - f \| \mathrm{d}t = \int_0^\infty e^{-s} \| T_{\frac{s}{\lambda}} f - f \| \mathrm{d}s \to 0.$$

由 (37.7) 可知 R_λ 的值域 $\mathfrak{R}_\lambda$ 与 λ 无关. 实际上,

$$R_\mu f = R_\lambda[I + (\lambda - \mu)R_\mu]f \in \mathfrak{R}_\lambda,$$

故 $\mathfrak{R}_\mu \subset \mathfrak{R}_\lambda$. 将 λ, μ 互换, 则得 $\mathfrak{R}_\lambda \subset \mathfrak{R}_\mu$, 故 $\mathfrak{R}_\lambda = \mathfrak{R}_\mu$. 因此 $\mathfrak{R}_\lambda$ 可简写为 $\mathfrak{R}$.

$\mathfrak{R}_\lambda$ 显然为线性子空间, 由 (37.8) 得

$$\overline{\mathfrak{R}}_\lambda = E \quad (\text{“—” 表示闭包}), \tag{37.9}$$

即 $\mathfrak{R}_\lambda$ 为 E 的稠密线性子空间.

其次, 试证 $\mathfrak{D}(A) = \mathfrak{R}$. 先设 $f \in \mathfrak{R}$. 因 $f = R_\lambda g, g \in E$, 故

$$\begin{aligned}
T_s f = T_s R_\lambda g &= \int_0^\infty e^{-\lambda t} T_{t+s} g \mathrm{d}t \\
&= e^{\lambda s}\int_s^\infty e^{-\lambda t} T_t g \mathrm{d}t \\
&= e^{\lambda s}\int_0^\infty e^{-\lambda t} T_t g \mathrm{d}t - e^{\lambda s}\int_0^s e^{-\lambda t} T_t g \mathrm{d}t,
\end{aligned}$$

$$\frac{T_s f - f}{s} = \frac{e^{\lambda s}-1}{s} R_\lambda g - e^{\lambda s}\frac{1}{s}\int_0^s e^{-\lambda t} T_t g \mathrm{d}t \to \lambda f - g.$$

故若 $f \in \mathfrak{R}(f = R_\lambda g)$, 则 $f \in \mathfrak{D}(A)$, 且

$$Af = \lambda f - g.$$

注意: $\lambda - A$ 是一一对应的, 为证明这一点, 只需由 $(\lambda - A)f = 0$ 能导出 $f = 0$ 即可. 由 $(\lambda - A)f = 0$ 得 $A_f = \lambda f$, 故

$$T_s f = f + \lambda s f + o(s) = (1 + \lambda s) f + o(s),$$

$$\| f \| \geqslant \| T_s f \| = (1 + \lambda s) \| f \| + o(s),$$

$$0 \geqslant \lambda s \| f \| + o(s), \quad 故 \quad \| f \| + o(1) \leqslant 0.$$

由此即得 $\| f \| \leqslant 0$, 故 $\| f \| = 0$.

现在反过来证明, 如果 $f \in \mathfrak{D}(A)$, 则 $f \in \mathfrak{R}$. 于是就得证 $\mathfrak{D}(A) = \mathfrak{R}$. 由于 $f \in \mathfrak{D}(A)$, 故 Af 存在, 设 $g = \lambda f - Af$, 又设 $f_0 = R_\lambda g$. 由上半证明之末所述有

$$Af_0 = \lambda f_0 - g,$$

故

$$(\lambda - A) f_0 = g = (\lambda - A) f.$$

因 $\lambda - A$ 为一对一的, 故

$$f = f_0 = R_\lambda g \in \mathfrak{R}.$$

总结上述诸结果, 可知 $\mathfrak{D}(A) = \mathfrak{R}$, 且如果 $f \in \mathfrak{D}(A)$(因之 $f = R_\lambda g$), 则

$$Af = \lambda f - g. \tag{37.10}$$

$\lambda - A$ 为一对一的, 因此

$$(\lambda - A)^{-1} = R_\lambda, \ 因而 \ R_\lambda^{-1} = \lambda - A, \tag{37.11}$$

而且

$$\overline{\mathfrak{D}(A)} = \overline{\mathfrak{R}} = E.$$

若 $f \in \mathfrak{D}(A)$, 则 $T_t f \in \mathfrak{D}(A)$, 且下面的生成方程 (evolution equation) 成立:

$$\frac{\mathrm{d}T_t f}{\mathrm{d}t} = AT_t f (= T_t A f). \tag{37.12}$$

因为, 如 $f \in \mathfrak{D}(A)$, 则 $f = R_\lambda g$, 故

$$\begin{aligned}T_tf&=T_tR_\lambda g=R_\lambda T_tg\in\mathfrak{D}(A),\\ \frac{\mathrm{d}T_tf}{\mathrm{d}t}&=\lim_{\delta\downarrow 0}\frac{T_\delta-I}{\delta}T_tf\\ &=AT_tf=AT_tR_\lambda g\\ &=AR_\lambda T_tg=\lambda R_\lambda T_tg-T_tg\\ &=T_t(\lambda R_\lambda g-g)=T_tAR_\lambda g=T_tAf.\end{aligned}$$

§38 Hille-Yosida 理论 (2) 半群的构造

上节中设半群已知而叙述了其生成算子的性质, 以及它与半群的关系. 设 A 为生成算子, 则显然有

(A.1) A 为线性算子, 且 $\overline{\mathfrak{D}(A)}=E$.

(A.2) $(\lambda-A)^{-1}$ 定义于整个 E 上, 且

$$\|(\lambda-A)^{-1}\|\leqslant\frac{1}{\lambda}\quad(\lambda>0).$$

本节的目的正与此相反, 即要证明, 满足上述两条件的 A 是某唯一半群的生成算子.

首先, 试由上列条件证明:

$$\text{若}I_\lambda=\lambda(\lambda-A)^{-1},\quad\text{则}\|I_\lambda f-f\|\to 0\ (\lambda\to\infty).\tag{38.1}$$

如 $f\in\mathfrak{D}(A)$, 则

$$\begin{aligned}I_\lambda f-f&=(\lambda-A)^{-1}\lambda f-(\lambda-A)^{-1}(\lambda-A)f\\ &=(\lambda-A)^{-1}(\lambda f-(\lambda-A)f)=(\lambda-A)^{-1}Af,\end{aligned}$$

$$\|I_\lambda f-f\|\leqslant\|Af\|/\lambda\to 0\quad(\lambda\to\infty).$$

对一般的 $f\in E$, 有

$$f=g+h,\qquad g\in\mathfrak{D}(A),\quad\|h\|<\varepsilon.$$

由上述知 $\|I_\lambda g-g\|\to 0$.

$$\begin{aligned}\|I_\lambda f-f\|&\leqslant\|I_\lambda g-g\|+\|I_\lambda h\|+\|h\|\\ &\leqslant\|I_\lambda g-g\|+2\|h\|<\|I_\lambda g-g\|+2\varepsilon\to 2\varepsilon.\end{aligned}$$

故对一般的 f 也有 $\| I_\lambda f - f \| \to 0$.

由上节, 如 A 有界, 则 $T_t = \mathrm{e}^{tA}$ 是以 A 为生成算子的半群. 但一般说来, 因 A 不是有界的, 故不能简单地得出这样的结论. 现在先用 $A_\lambda = AI_\lambda$ 来代替 A. 如上所述, I_λ 可看成恒等算子 I 的近似算子, 故 A_λ , 亦可看成 A 的近似算子, 且由

$$\begin{aligned} A_\lambda &= A \cdot \lambda \cdot (\lambda - A)^{-1} = (\lambda - (\lambda - A))\lambda(\lambda - A)^{-1} \\ &= \lambda \cdot \lambda(\lambda - A)^{-1} - \lambda I = \lambda(I_\lambda - I) \end{aligned}$$

可知 A_λ 有界. 今设

$$T_t^{(\lambda)} = \mathrm{e}^{tA_\lambda}, \tag{38.2}$$

这就是所求的 T_t 的近似半群. 试证 $T_t^{(\lambda)}$ 为半群. 显然 $T_t^{(\lambda)} \to I, T_t^{(\lambda)} \cdot T_s^{(\lambda)} = T_{t+s}^{(\lambda)}$. $\| T_t^{(\lambda)} \| \leqslant 1$ 的证明如下:

$$\begin{aligned} \| T_t^{(\lambda)} \| &= \| \mathrm{e}^{tA_\lambda} \| = \| \mathrm{e}^{t\lambda(I_\lambda - I)} \| = \| \mathrm{e}^{-t\lambda}\mathrm{e}^{t\lambda I_\lambda} \| \\ &\leqslant \mathrm{e}^{-t\lambda} \| \mathrm{e}^{t\lambda I_\lambda} \| \leqslant \mathrm{e}^{-t\lambda} \sum_n \frac{\| t\lambda I_n \|^n}{n!} \leqslant \mathrm{e}^{-t\lambda} \sum \frac{(t\lambda)^n}{n!} = 1. \end{aligned}$$

其次, 令

$$T_t = \lim_{\lambda \to \infty} T_t^{(\lambda)} \tag{38.3}$$

来求 T_t, 为此, 有必要证明其收敛性.

首先注意

$$(\lambda - A)^{-1}(\mu - A)^{-1} = (\mu - A)^{-1}(\lambda - A)^{-1}. \tag{38.4}$$

为此, 只需证

$$(\lambda - A)(\mu - A)(\lambda - A)^{-1}(\mu - A)^{-1} = I.$$

由

$$\mu - A = (\mu - \lambda)I + (\lambda - A), \quad \lambda - A = (\lambda - \mu)I + (\mu - A),$$

再注意 $(\lambda - A)^{-1}$ 和 $(\mu - A)^{-1}$ 的值域分别落入 $\lambda - A$ 和 $\mu - A$ 的定义域内, 故只要作些形式的计算即得证. 由 (38.4) 直接可得

$$I_\lambda I_\mu = I_\mu I_\lambda, \text{ 即 } A_\lambda A_\mu = A_\mu A_\lambda. \tag{38.5}$$

(回忆 $A_\lambda=\lambda(I_\lambda-I)$.) 由此可知 $T_t^{(\lambda)}$,$T_t^{(\mu)}$ 能与 $A_\lambda, A\mu$ 中任一个互相交换. 又由 $T_t^{(\lambda)}=\mathrm{e}^{tA_\lambda}$ 得

$$\frac{\mathrm{d}T_t^{(\lambda)}f}{\mathrm{d}t}=A_\lambda T_t^{(\lambda)}f,\quad \frac{\mathrm{d}T_t^{(\mu)}f}{\mathrm{d}t}=A_\mu T_t^{(\mu)}f.$$

故若设 $f_t=T_t^{(\lambda)}f-T_t^{(\mu)}f$, 则得

$$\frac{\mathrm{d}f_t}{\mathrm{d}t}=A_\lambda T_t^{(\lambda)}f-A_\mu T_t^{(\mu)}f=A_\lambda\cdot f_t+g_t, \tag{38.6}$$

其中

$$g_t=(A_\lambda-A_\mu)T_t^{(\mu)}f=T_t^{(\mu)}(A_\lambda f-A_\mu f). \tag{38.7}$$

用解普通微分方程同样的方法解 (38.6), 得

$$\begin{aligned}f_t&=\int_0^t \mathrm{e}^{(t-s)A_\lambda}g_s\mathrm{d}s+f_0=\int_0^t \mathrm{e}^{(t-s)A_\lambda}g_s\mathrm{d}s\\&=\int_0^t T_{t-s}^{(\lambda)}g_s\mathrm{d}s.\end{aligned}$$

故 $\|f_t\|\leqslant\int_0^t\|T_{t-s}^{(\lambda)}g_s\|\mathrm{d}s\leqslant\int_0^t\|g_s\|\mathrm{d}s\leqslant\|A_\lambda f-A_\mu f\|\cdot t$. 因此, 若 $f\in\mathfrak{D}(A)$, 则 $A_\lambda f-A_\mu f=(I_\lambda-I_\mu)Af\to 0(\lambda,\mu\to\infty)$. 故当 $\lambda,\mu\to\infty$ 时, 对有界区间内的 t, $f_t=T_t^{(\lambda)}f-T_t^{(\mu)}f$ 一致地趋于 0(广义的一致收敛).

至于对任意的 f, 则取与 f 接近的 $g\in\mathfrak{D}(A)$, 并注意

$$\|T_t^{(\lambda)}f-T_t^{(\mu)}f\|\leqslant 2\|f-g\|+\|T_t^{(\lambda)}g-T_t^{(\mu)}g\|,$$

可见 $T_t^{(\lambda)}f$ 对于 t 仍为广义的一致收敛. 设其极限为 T_tf, 易证 T_t 为 E 上的半群. 今证 T_t 的生成算子 $\widetilde{A}$ 与原来的 A 相等.

对于 $f\in\mathfrak{D}(A)$, 有

$$T_t^{(\lambda)}f-f=\int_0^t T_s^{(\lambda)}A_\lambda f\mathrm{d}s=\int_0^t T_s^{(\lambda)}I_\lambda Af\mathrm{d}s.$$

由于

$$\begin{aligned}\|T_s^{(\lambda)}I_\lambda Af-T_sAf\|&\leqslant\|T_s^{(\lambda)}I_\lambda Af-T_s^{(\lambda)}Af\|+\|T_sAf-T_s^{(\lambda)}Af\|\\&\leqslant\|I_\lambda Af-Af\|+\|T_sAf-T_s^{(\lambda)}Af\|\to 0\end{aligned}$$

(在 $0\leqslant s\leqslant t$ 内为一致收敛),

故

$$T_t f - f = \int_0^t T_s A f \mathrm{d}s.$$

由此得

$$\frac{T_t f - f}{t} \to Af, \text{ 即 } f \in \mathfrak{D}(\widetilde{A}), \text{且 } \widetilde{A}f = Af.$$

这表示 $\widetilde{A} \supset A$. 由上节所述, $(\lambda - \widetilde{A})^{-1}$ 存在, 又由关于 A 的假定, 可知 $(\lambda - A)^{-1}$ 也存在, 故由 $\widetilde{A} \supset A$ 可得 $(\lambda - \widetilde{A})^{-1} \supset (\lambda - A)^{-1}$. 又因 $(\lambda - A)^{-1}$ 为定义在整个 E 上的算子, 故 $(\lambda - \widetilde{A})^{-1} = (\lambda - A)^{-1}$, 故得 $\widetilde{A} = A$.

以 A 为生成算子的半群的存在性已经证明, 为证其唯一性, 只要证明, 若存在另一如此的半群 S_t, 则 $T_t = S_t$. 由于 $\mathfrak{D}(A)$ 在 E 内稠密, T_t, S_t 有界, 故只要对 $f \in \mathfrak{D}(A)$, 证明 $T_t f = S_t f$ 即可.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(S_t f - T_t^{(\lambda)} f) &= A S_t f - A_\lambda T_t^{(\lambda)} f \\ &= A_\lambda (S_t f - T_t^{(\lambda)} f) + S_t (I - I_\lambda) A f.\end{aligned}$$

与解 (38.6) 同样解上式得

$$S_t f - T_t^{(\lambda)} f = \int_0^t T_{t-u}^{(\lambda)} S_u (I - I_\lambda) A f \mathrm{d}u.$$

故

$$\| S_t f - T_t^{(\lambda)} f \| \leqslant t \| (I - I_\lambda) A f \| \to 0, \quad \lambda \to \infty,$$

由此得

$$S_t f = T_t f.$$

注意 由上述证明, 可见条件 (A.2) 对所有的 $\lambda > 0$ 成立并非必要. 如对 $\lambda = \lambda_n \to \infty$, 能使 (A.2) 满足, 则以 A 作为生成算子的半群 T_t 唯一存在.

§39 转移概率的生成算子 (1) 一般理论

设 $P(t, x, E)$ 为 §35 中所引入的转移概率. 如 §36 所述, 此转移概率对应于 C 上的半群 $T_t (t > 0)$. $T_t (t > 0)$ 的生成算子 A 称为转移概率

$P(t,x,E)$ 的生成算子. 若 $f \in \mathfrak{D}(A)$, 则

$$Af(x) = \lim_{t \downarrow 0} \frac{\int_R f(y)P(t,x,\mathrm{d}y) - f(x)}{t}. \tag{39.1}$$

下列三条件是等价的.

(i) $f \in \mathfrak{D}(A)$,

(ii) $\varphi_t(x) \equiv \frac{1}{t}\left[\int_R f(y)P(t,x,\mathrm{d}y) - f(x)\right]$ 当 $t \to 0$ 时, 关于 x 一致收敛,

(iii) $\varphi_t(x)$ 在 $t \to 0$ 时, 对每一 x 都收敛, 且其极限关于 x 连续.

(i) $\to$(ii), (i)$\to$(iii) 显然. 因 $\varphi(x)$ 是 x 的连续函数, 若 (ii) 成立, 则 $\varphi_t(x)$ 的极限 $\varphi(x)$ 也连续, 并且 $\| \varphi_t - \varphi \| \to 0$, 故 $f \in \mathfrak{D}(A)$. 今试由 (iii) 导出 (i). 把使 (iii) 成立的 f 全体记为 $\widetilde{\mathfrak{D}}$, 对 $f \in \widetilde{\mathfrak{D}}$, 定义

$$\widetilde{A}f(x) = \lim_{t \downarrow 0} \varphi_t(x).$$

显然 $\widetilde{A}f \in \boldsymbol{C}$. 如 $f \in \mathfrak{D}(A)$, 则 $f \in \widetilde{\mathfrak{D}}$, 且 $Af = \widetilde{A}f$. 故 $A \subset \widetilde{A}$. 注意 $\lambda - \widetilde{A}$ 为一一对应的, 为证此, 先注意若 $f(x)$ 在 $x = x_0$ 处取最小值, 则

$$\widetilde{A}f(x_0) = \lim_{t \downarrow 0} \varphi_t(x_0) \geqslant 0.$$

为了证明 $\lambda - \widetilde{A}$ 为一一对应的, 只要证明

$$(\lambda - \widetilde{A})u = 0 \to u = 0.$$

因 u 为连续函数, 故有最小值 $u(x_0)$.

$$u(x_0) = \frac{1}{\lambda}\widetilde{A}u(x_0) \geqslant 0, \text{ 故恒有 } u(x) \geqslant 0.$$

同样因 $(\lambda - \widetilde{A})(-u) = 0$, 故 $-u(x) \geqslant 0$, 于是 $u(x) \equiv 0$.

由 $A \subset \widetilde{A}$ 得 $\lambda - A \subset \lambda - \widetilde{A}$. 因 $\lambda - A$ 有逆算子 $(\lambda - A)^{-1} (= R_\lambda)$, 并由上述知 $(\lambda - \widetilde{A})^{-1}$ 也存在, 故 $R_\lambda = (\lambda - A)^{-1} \subset (\lambda - \widetilde{A})^{-1}$. 因 R_λ 的定义域为全空间 $\boldsymbol{C}$, 故得

$$(\lambda - \widetilde{A})^{-1} = R_\lambda = (\lambda - A)^{-1}, \quad \text{即 } \widetilde{A} = A.$$

故 $\widetilde{\mathfrak{D}} = \mathfrak{D}(A)$, 因此 (iii)$\to$(i).

上节中已证明 Banach 空间的算子 A 产生半群的充要条件为 (A.1) 及 (A.2). 下面将求特别是 A 为转移概率 $P(t,x,E)$ 的生成算子的条件.

定理 1 A 为转移概率的生成算子的充要条件是:

(a.1) A 是 $\boldsymbol{C}$ 上的线性算子, 且 $\overline{\mathfrak{D}(A)}=\boldsymbol{C}$,

(a.2) $A\cdot 1=0$,

(a.3) 若 $u\in\boldsymbol{C}$ 在 $x=x_0$ 处取最小值, 则 $Au(x_0)\geqslant 0$,

(a.4) 对于 $\lambda>0, (\lambda-A)u=v$ 对一切 $v\in\boldsymbol{C}$ 必有解 u.

证明 必要性易由定义及半群的一般理论推出, 今只证其充分性. 先证 $(\lambda-A)^{-1}$ 是定义在全 $\boldsymbol{C}$ 上的. 由 (a.4) 可知, 只需证明 $\lambda-A$ 为一一对应, 即由 $\lambda u=Au$ 能得 $u=0$ 即可. 因 $u(x)$ 为连续函数, 故取最小值 $u(x_0)$,

$$u(x_0)=\frac{1}{\lambda}Au(x_0)\geqslant 0 \quad (\text{由 (a.3)}).$$

故恒有 $u(x)\geqslant 0$. 又因 $\lambda(-u)=A(-u)$, 利用同样的论证得 $-u(x)\geqslant 0$, 故 $u(x)\equiv 0$.

其次证 $\|(\lambda-A)^{-1}\|\leqslant\frac{1}{\lambda}$. 设 $u=(\lambda-A)^{-1}v$, 令 $u(x)$ 之最小值为 $u(x_0)$, 则

$$v(x_0)=(\lambda-A)u(x_0)=\lambda u(x_0)-Au(x_0)\leqslant\lambda u(x_0).$$

故

$$u(x)\geqslant u(x_0)\geqslant\frac{1}{\lambda}v(x_0)\geqslant-\frac{1}{\lambda}\|v\|.$$

又因 $(-u)=(\lambda-A)^{-1}(-v)$, 故

$$-u(x)\geqslant-\frac{1}{\lambda}\|-v\|=-\frac{1}{\lambda}\|v\|,$$

即

$$u(x)\leqslant\frac{1}{\lambda}\|v\|,$$

由此即得

$$\|u\|\leqslant\frac{1}{\lambda}\|v\|.$$

于是上节的 (A.1)、(A.2) 均已满足, 故 A 为 $\boldsymbol{C}$ 上某半群 $T_t(t>0)$ 的生成算子.

试证 $T_t \geqslant 0$, 亦即如 $u \geqslant 0$(即对一切 $x, u(x) \geqslant 0$), 则 $T_t u \geqslant 0$. 为此先证 $(\lambda - A)^{-1} \geqslant 0$, 也就是说, 设 $(\lambda - A)u = v$, 而 $v \geqslant 0$, 则 $u \geqslant 0$. 令 $u(x)$ 的最小值为 $u(x_0)$, 则 $Au(x_0) \geqslant 0$, 故

$$\lambda u(x_0) = Au(x_0) + v(x_0) \geqslant 0,$$

亦即

$$u \geqslant 0.$$

故得 $I_\lambda = \lambda(\lambda - A)^{-1} \geqslant 0$. 由此得

$$T_t^{(\lambda)} = \mathrm{e}^{tA_\lambda} = \mathrm{e}^{t\lambda(I_\lambda - I)} = \mathrm{e}^{-t\lambda}\mathrm{e}^{tI_\lambda} = \mathrm{e}^{-t\lambda}\sum \frac{(tI_\lambda)^n}{n!} \geqslant 0,$$

故

$$T_t = \lim_{\lambda \to \infty} T_t^{(\lambda)} \geqslant 0.$$

其次, 试证 $T_t 1 = 1$. 因 $(\lambda - A)\frac{1}{\lambda} = 1$(参看 (a.2)), 故

$$\frac{1}{\lambda} = (\lambda - A)^{-1} 1 = \int_0^\infty \mathrm{e}^{-\lambda t} T_t \cdot 1 \mathrm{d}t,$$

故得 $T_t 1 = 1$.

因 $T_t f(x)$ 对每一 t, x 而言皆为 f 的线性泛函数, 而且

$$\text{若 } f \geqslant 0, \text{ 则 } T_t f(x) \geqslant 0,$$
$$T_t 1(x) = 1,$$

故由 Riesz 定理, 可知存在概率分布 $P(t, x, E)$, 使

$$T_t f(x) = \int_R f(y) P(t, x, \mathrm{d}y).$$

$P(t, x, E)$ 满足转移概率诸条件, 可由 $T_t (t > 0)$ 是半群而容易导出.

注意　与上节 (A.2) 同样, (a.4) 也只需对 $\lambda = \lambda_n \to \infty$ 成立即可.

§40　转移概率的生成算子 (2) 例题

例 1　上面已看到, R 为有限集时, 转移概率对应于一个矩阵的半群 P_t. 这时 C 为向量空间, 其维数等于 R 中元素的个数. 若将其中向量写

成列向量, 则 $T_t f$ 等于 $P_t f$, 因此矩阵的半群 $P_t(t>0)$ 可以看成伴随转移概率的半群. 故生成算子 $\boldsymbol{A}$ 亦可看成矩阵, 且

$$\boldsymbol{A}f=\lim_{t\downarrow 0}\frac{T_t-I}{t}f,\quad f\in\mathfrak{D}(\boldsymbol{A}).$$

此时由 $\overline{\mathfrak{D}(\boldsymbol{A})}=\boldsymbol{C}$ 可得 $\mathfrak{D}(\boldsymbol{A})=\boldsymbol{C}$(因 $\boldsymbol{C}$ 为有限维), 故上式对任意 f 均成立, 且

$$\boldsymbol{A}=\lim_{t\downarrow 0}\frac{T_t-I}{t}.$$

因 $T_t\geqslant 0$(此表示 T_t 的诸元素均 $\geqslant 0$), 故若令 $\boldsymbol{A}=(a_{ij})$, 则

$$a_{ij}\geqslant 0\quad(i\neq j).\tag{40.1}$$

又因 $T_t\mathbf{1}=\mathbf{1}$($\mathbf{1}$ 为所有元素均为 1 的向量), 故 $\boldsymbol{A}\mathbf{1}=0$, 因此

$$\sum_j a_{ij}=0.\tag{40.2}$$

此二条件为必要, 但亦为充分. 为此, 试验证上节的 (a.1)~(a.4) 成立.(a.1), (a.2) 是显然的. 设向量 $\boldsymbol{u}$ 的元素中最小者为 u_{i_0}.

$$\begin{aligned}(\boldsymbol{A}\boldsymbol{u})_{i_0}&=\sum_j a_{i_0 j}u_j=a_{i_0 i_0}u_{i_0}+\sum_{j\neq i_0}a_{i_0 j}u_j\\&\geqslant a_{i_0 i_0}u_{i_0}+\sum_{j\neq i_0}a_{i_0 j}u_{i_0}\qquad(\text{由 }(40.1))\\&=0.\qquad\qquad\qquad\qquad\qquad(\text{由 }(40.2))\end{aligned}$$

此即表示 (a.3) 成立. 现证明对充分大的 λ,

$$(\lambda-\boldsymbol{A})\boldsymbol{u}=\boldsymbol{v}\tag{40.3}$$

有解. 形式地得

$$\boldsymbol{u}=(\lambda-\boldsymbol{A})^{-1}\boldsymbol{v}=\lambda^{-1}\left(1-\frac{\boldsymbol{A}}{\lambda}\right)^{-1}\boldsymbol{v}=\lambda^{-1}\left[1+\left(\frac{\boldsymbol{A}}{\lambda}\right)+\left(\frac{\boldsymbol{A}}{\lambda}\right)^2+\cdots\right]\boldsymbol{v},$$

为使上式收敛, 只要令 $\lambda>\|\boldsymbol{A}\|$. 事实上, 对此种 λ 可知上式右边确定而且给出 (40.3) 的解.

例 2 试求 §35 例 2 中的转移概率的生成算子. 为此要利用 §37 中的

$$\mathfrak{D}(A)=\mathfrak{R},\quad AR_\lambda f=\lambda R_\lambda f-f.$$

以后用 $x, y, \cdots$ 表有限数, 因

$$T_t f(x) = \int N_t(y-x) f(y) \mathrm{d}y, \quad T_t f(\infty) = f(\infty),$$

故

$$R_\lambda f(x) = \int R_\lambda(y-x) f(y) \mathrm{d}y, \quad R_\lambda f(\infty) = \frac{1}{\lambda} f(\infty).$$

这里

$$R_\lambda(x) = \int_0^\infty \mathrm{e}^{-\lambda t} N_t(x) \mathrm{d}t = \frac{1}{\sqrt{2\lambda}} \mathrm{e}^{-\sqrt{2\lambda}|x|},$$

故

$$R_\lambda f(x) = \int \frac{1}{\sqrt{2\lambda}} \mathrm{e}^{-\sqrt{2\lambda}|y-x|} f(y) \mathrm{d}y.$$

由上式可知 $R_\lambda f(x)$ 为二次连续可微, 且

$$(R_\lambda f)''(x) = 2\lambda R_\lambda f(x) - 2f(x),$$

故 $\lim\limits_{x\to\infty} (R_\lambda f)''(x) = 0$, 由此即得

$$\mathfrak{D}(A) \subset \widetilde{\mathfrak{D}} \equiv \{g/g \in \boldsymbol{C}, g''(x) \text{ 连续}, \lim_{x\to\infty} g''(x) = 0\},$$

且对于 $g(= R_\lambda f) \in \mathfrak{D}(A)$ 可得

$$Ag(x) = \lambda g(x) - f(x) = \frac{1}{2} g''(x), \qquad Ag(\infty) = 0.$$

今考虑使 $\mathfrak{D}(\widetilde{A}) = \widetilde{\mathfrak{D}}, \widetilde{A}g(x) = \frac{1}{2}g''(x), \widetilde{A}g(\infty) = 0$ 成立的 $\widetilde{A}$. 显然 $\widetilde{A} \supset A$. 次证 $(\lambda - \widetilde{A})^{-1}$ 的存在性. 为此要由 $\lambda u = \widetilde{A}u$ 推出 $u = 0$. $\lambda u = \widetilde{A}u$ 的解 u 满足

$$\lambda u(x) = \frac{1}{2} u''(x),$$

即

$$u(x) = a\mathrm{e}^{\sqrt{2\lambda}x} + b\mathrm{e}^{-\sqrt{2\lambda}x}.$$

因 $u(x)$ 有界, 故得 $a = b = 0$, 亦即 $u = 0$, 因此,

$$(\lambda - \widetilde{A})^{-1} \supset (\lambda - A)^{-1} = R_\lambda,$$

故

$$(\lambda - \widetilde{A})^{-1} = (\lambda - A)^{-1},$$

亦即 $\widetilde{A}=A$.

例 3 考虑 §35 中例 3, 利用与上例相同的论证, 可知 $\mathfrak{D}(A)$ 的元素 f 可由下列诸条件来表达:

$$f \in \boldsymbol{C}, f''(x) \text{ 连续 } (0<x<\infty),$$
$$\lim_{x\to 0} f''(x) \text{ 存在, 有限}, \quad \lim_{x\to\infty} f''(x)=0.$$
$$\lim_{x\to 0} f'(x)=0.$$

且

$$Af(x)=\frac{1}{2}f''(x), \quad Af(0)=\frac{1}{2}f''(0+), \quad Af(\infty)=0.$$

例 4 设 R 为以 mod 1 来考虑的实数系, $\boldsymbol{C}$ 为定义于实数系上周期为 1 的连续函数空间, 以 $\mathfrak{D}$ 表 $\boldsymbol{C}$ 中之二次连续可微函数全体, 且对 $f\in\mathfrak{D}$, 定义

$$Af=f'.$$

试证 A 为某转移概率之生成算子, 为此必须验证上节 (a.1)~(a.4) 满足. (a.1)~(a.3) 是显然的. 为证 (a.4), 只要解方程

$$\lambda u-u'=v$$

即可, 此处 v 是周期为 1 的函数, u 必须自这类函数中求出.

$$\lambda \mathrm{e}^{-\lambda x}u(x)-\mathrm{e}^{-\lambda x}u'(x)=\mathrm{e}^{-\lambda x}v(x),$$

$$(-\mathrm{e}^{-\lambda x}u(x))'=\mathrm{e}^{-\lambda x}v(x).$$

因当 $x=+\infty$ 时, $\mathrm{e}^{-\lambda x}u(x)$ 为 0, 故

$$\mathrm{e}^{-\lambda x}u(x)=\int_x^\infty \mathrm{e}^{-\lambda y}v(y)\mathrm{d}y,$$

$$u(x)=\int_x^\infty \mathrm{e}^{-\lambda(y-x)}v(y)\mathrm{d}y.$$

此解确能满足上面的方程, 且由

$$u(x+1)=\int_{x+1}^\infty \mathrm{e}^{-\lambda(y-x-1)}v(y)\mathrm{d}y=\int_x^\infty \mathrm{e}^{-\lambda(y-x)}v(y+1)\mathrm{d}y$$
$$=u(x) \quad (\text{注意 } v(y+1)=v(y)),$$

故 $u(x)$ 有周期 1.

由上述得

$$R_\lambda v(x) = \int_x^\infty \mathrm{e}^{-\lambda(y-x)} v(y)\mathrm{d}y = \int_0^\infty \mathrm{e}^{-\lambda t} v(x+t)\mathrm{d}t.$$

故 $T_t v(x) = v(x+t)$, 因此

$$P(t,x,E) = \delta(x+t,E).$$

显然 $x+t$ 亦是 mod 1 的实数.

例 5 设 R 与上例同样为 mod 1 的实数系, 且 $Au = u''/2$. 因此 A 亦满足 (a.1)~(a.4), 故 A 为某 $P(t,x,E)$ 的生成算子. 设 $\mathfrak{D}(A)$ 为二次连续可微函数全体, $P(t,x,E)$ 由下式给出:

$$P(t,x,E) = \int_E \widetilde{N}_t(y-x)\mathrm{d}y,$$

$$\widetilde{N}_t(x) = \sum_{n=-\infty}^{\infty} N_t(x+2n) = \frac{1}{\sqrt{2\pi t}} \sum_{n=-\infty}^{\infty} \mathrm{e}^{-\frac{(x+2n)^2}{2t}}.$$

例 6 R 与上例中一样. 令 $Au = u'''$, 又令 $\mathfrak{D}(A)$ 为三次连续可微函数全体, 则此时不对应于转移概率, 因为 (a.3) 不成立. 为证 (a.3) 不成立, 只需找出三次连续可微而周期为 1 的函数 u, 它在零点取最小值, 而且 $u'''(0) < 0$. 这只要取周期为 1 的函数 $u(x)$, 使 $u(x) \geqslant 0$, 并在 $x = 0$ 处的展开为

$$u(x) = ax^2 + bx^3 + \cdots (a > 0, b < 0)$$

即可. 例如, 下面的函数就是此类函数:

$$u(x) = 2(\sin 2\pi x)^2 - (\sin 2\pi x)^3.$$

§41 Markoff 过程 (1) Markoff 性

设 $P(t,x,E)$ 是 R 上的转移概率. R 及 P 应满足 §35 中的条件. 转移概率 $P(t,x,E)$ 的直观意义是: 最初处于状态 x 经过时间 t 后转入 E 中的状态的概率. 亦即, 它给出了运动体系随时间变化的概率法则. 今以

$x(t)$ 表此体系在时刻 t 所处的状态. $x(t)$ 显然是随机过程, 从测度论的观点看来, 应该写成 $x(t,\omega),\omega\in\Omega(\boldsymbol{B},P)$. 若设在 $t=0$ 时的概率分布, 即初始状态 $x(0,\omega)$ 的概率分布为 $\varPhi$, 则 $x(t,\omega)$ 的有限维联合分布显然为

$$\begin{aligned}&P(x(0,\omega)\in E_0,x(t_1,\omega)\in E_1,x(t_2,\omega)\in E_2,\cdots,x(t_n,\omega)\in E_n)\\&=\int_{E_0}\varPhi(\mathrm{d}\xi_1)\int_{E_1}P(t_1,\xi_1,\mathrm{d}\xi_2)\int_{E_2}P(t_2-t_1,\xi_2,\mathrm{d}\xi_3)\cdots\\&\quad\int_{E_n}P(t_n-t_{n-1},\xi_{n-1},\mathrm{d}\xi_n).\end{aligned}\tag{41.1}$$

这样的随机过程 $x(t,\omega)$ 是否存在? 设 $\Omega=R^{[0,\infty)},\omega=\prod_t\omega_t,x(t,\omega)=\omega_t$, 则问题化为能否在 Ω 上引进概率 P, 使 (41.1) 成立. 这个可能性与实数情形 (Kolmogoroff 定理) 一样可以证明. 因为这样的 $x(t,\omega)$ 本质上只有一个, 故称它为具有转移概率 $P(t,x,E)$ 及初始概率 $\varPhi$ 的随机过程. 这类过程最初由 Markoff 所研究, 故也称为 **Markoff 过程** (Markoff process). $x(t,\omega)$ 的概率法则 (有限维分布) 虽与初始概率 $\varPhi$ 有关, 但只要讨论 $\varPhi$ 为 δ 分布 $\delta(a,E)$ 的情形, 则一般情形可由 $\varPhi(\mathrm{d}a)$ 积分而得. $\delta(a,E)$ 的情形就是 $x(0,\omega)\equiv a$, 亦即从状态 a 出发的情形. 这样的 Markoff 过程记为 $x^{(a)}(t,\omega)$. 本来 ω 也应写成 $\omega^{(a)}$, 但从前后关系看来容易明白, 故省略. 以后称 Markoff 过程族 $x^{(a)}(t,\omega),a\in R$, 为对应于 $P(t,x,E)$ 的 Markoff 过程. 概率论的目的是研究 $x^{(a)}(t,\omega)$, 迄今有关转移概率及其半群的叙述是为此目的所作的准备.

Markoff 过程 $x^{(a)}(t,\omega)$ 的概率性质中, 特别显著的是 Markoff 性.

(i) 若以 $\boldsymbol{B}_t$(必要时亦写为 $\boldsymbol{B}_t^{(a)}$) 表示包含如下形状

$$\{\omega/x^{(a)}(s,\omega)\in E\},s\leqslant t,E\in\boldsymbol{B}_R$$

的 ω 集的最小 Borel 体, 则以概率为 1 地有

$$\begin{aligned}P(x^{(a)}(t+u,\omega)\in E/\boldsymbol{B}_t)&=P(x^{(b)}(u)\in E)_{b=x^{(a)}(t,\omega)}\\&=P(u,x^{(a)}(t,\omega),E).\end{aligned}\tag{41.2}$$

直观地说, 若在 t 时的状态确定, 则以后运动的状态与 t 以前的无关, 好像从 t 时的状态出发一样, 这种性质称为 Markoff 性. 要证明上面的关系式, 只需证明, 若 $M\in\boldsymbol{B}_t$, 则

$$P\left(\{\omega/x^{(a)}(t+u,\omega)\in E\}\cap M\right)=E\{P(u,x^{(a)}(t,\omega),E);M\}.\tag{41.2'}$$

当 M 为形如

$$\{\omega/x^{(a)}(t_1,\omega)\in E_1,\cdots,x^{(a)}(t_n,\omega)\in E_n\},\ 0<t_1<t_2<\cdots<t_n<t$$

的集时, 由 (41.1) 立得 (41.2′). 由于 (41.2′) 两边对 M 是可加的, 故易证对所有的 $M\in\boldsymbol{B}_t$ 均成立.

(ii) 上述 Markoff 性可稍推广如下: 如 $0<u_1<u_2<\cdots<u_m$, 则对 $M\in\boldsymbol{B}_t$ 以概率为 1 地有

$$\begin{aligned}&P\{x^{(a)}(t+u_1)\in E_1,x^{(a)}(t+u_2)\in E_2,\cdots,x^{(a)}(t+u_m)\in E_m/\boldsymbol{B}_t\}\\=&P\{x^{(b)}(u_1)\in E_1,x^{(b)}(u_2)\in E_2,\cdots,x^{(b)}(u_m)\in E_m\}_{b=x^{(a)}(t)}.\end{aligned}\tag{41.3}$$

这式可用与 (i) 同样的方法证明.

(iii) 再推广 (ii). 为此, 引进若干记号. $R^{[0,\infty)}$ 的意义如前所述. $\xi\in R^{[0,\infty)}$ 可以写成 $\prod\limits_{0\leqslant t<\infty}\xi_t$ 的形状, 而 ξ_t 称为 ξ 的 t 坐标, 并以 $p_t(\xi)$ 来表示. $\boldsymbol{B}(R^{[0,\infty)})$ 表示包含 $R^{[0,\infty)}$ 中形如 $P_t^{-1}(E)(E\in\boldsymbol{B}_R)$ 的子集的最小 Borel 体. 今设 $x^{(a)}(\cdot,\omega)$ 表示在 $\prod\limits_u x^{(a)}(u,\omega)$, 即在 $R^{[0,\infty)}$ 中取值的随机变量, 且其第 u 坐标等于 $x^{(a)}(t,\omega)$. 又设 $x^{(a)}(t+\cdot,\omega)$ 表示同样的随机变量, 其第 u 坐标等于 $x^{(a)}(t+u,\omega)$.

于是, 对于 $\varXi\in\boldsymbol{B}(R^{[0,\infty)})$, 以概率为 1 的有

$$P\{x^{(a)}(t+\cdot)\in\varXi/\boldsymbol{B}_t\}=P\{x^{(b)}(\cdot)\in\varXi\}_{b=x^{(a)}(t)}.\tag{41.4}$$

若 $\varXi$ 为形如 $P_{u_1}^{-1}(E_1)\cap P_{u_2}^{-1}(E_2)\cap\cdots\cap P_{u_m}^{-1}(E_m)$ 的集, 则上式即为 (ii) 中的 (41.3). 一般情况可利用两边对 $\varXi$ 的可加性证明之.

对应于上述诸性质, 可得关于条件期望的下述性质.

(iv) 如 f 为有界 $(\boldsymbol{B}_R)$ 可测函数, 则

$$E\{f(x^{(a)}(t+u))/\boldsymbol{B}_t\}=E\{f(x^{(b)}(u))\}_{b=x^{(a)}(t)}.\tag{41.5}$$

若 f 是可测集的示性函数, 则上式由 (i) 而显然成立. 对一般的 f, 可利用上式两边对 f 的可加性来证明.

(v) 如 f 是 $R(\boldsymbol{B}_R)$ 上的 n 元有界 $(\boldsymbol{B}_R)$ 可测函数, 则

$$\begin{aligned}&E\{f(x^{(a)}(t+u_1),x^{(a)}(t+u_2),\cdots,x^{(a)}(t+u_n))/\boldsymbol{B}_t\}\\=&E\{f(x^{(b)}(u_1),x^{(b)}(u_2),\cdots,x^{(b)}(u_n))\}_{b=x^{(a)}(t)}.\end{aligned}\tag{41.6}$$

(vi) 如 f 是 $R^{[0,\infty)}(\boldsymbol{B}(R^{[0,\infty)})$ 上的有界可测函数, 则

$$E\{f(x^{(a)}(t+\cdot))/\boldsymbol{B}_t\}=E\{f(x^{(b)}(\cdot))\}_{b=x^{(a)}(t)}. \tag{41.7}$$

与由 (i) 导出 (iv) 的方法一样, (v) 及 (vi) 可分别由 (ii) 及 (iii) 导出.

§42 Markoff 过程 (2) 样本过程的性质

在 §13 中, 已叙述过取实数值的随机过程的可分性. 但因现在所考虑的随机过程取值于具第二可数性的紧致空间 R, 故有必要重新定义可分性.

定义 在 R 内取值的随机过程 $x_t(t\in T)$ 称为**可分的**, 如果对于 R 上任意的实连续函数 $f, f(x_t)$ 是可分的.

设 $\{f_\lambda\}$ 为 R 上一族实连续函数, 若任意的连续函数能被 $\{f_\lambda\}$ 中有限个函数的线性组合来一致逼近时, 则 $\{f_\lambda\}$ 称为 R 上的实连续函数的**基** (base). 由于 R 是具第二可数性的紧致空间, 故知可数基存在.

上述定义中, 要求对所有的实连续函数, $f(x_t)$ 为可分, 但实际上只要求对于基中的函数就够了.

利用取实数值的随机过程的可分修正 (separable modification) 的存在, 可以证明取值于 R 的随机过程也有可分修正. 故不妨假定上节中定义的 Markoff 过程是可分的.

定理 1 对可分 Markoff 过程 $x^{(a)}(t)$ 的样本过程, "对于所有的 t, 两侧极限存在" 的概率为 1. 且对每个 t,

$$P(x(t+0)=x(t))=1. \tag{42.1}$$

证明 对 $f\in\boldsymbol{C}, f\geqslant 0$, 考虑

$$Y(t)=\mathrm{e}^{-\lambda t}R_\lambda f(x^{(a)}(t)).$$

$$\begin{aligned}
Y(t)&=\mathrm{e}^{-\lambda t}\int_0^\infty \mathrm{e}^{-\lambda s}T_s f(x^{(a)}(t))\mathrm{d}s\\
&=\mathrm{e}^{-\lambda t}\int_0^\infty \mathrm{e}^{-\lambda s}E(f(x^{(b)}(s)))_{b=x^{(a)}(t)}\mathrm{d}s\\
&=\mathrm{e}^{-\lambda t}\int_0^\infty \mathrm{e}^{-\lambda s}E\{f(x^{(a)}(t+s))/\boldsymbol{B}_t\}\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&= \int_t^\infty e^{-\lambda s} E\{f(x^{(a)}(s))/\boldsymbol{B}_t\}ds,\\
E\{Y(t+u)/\boldsymbol{B}_t\} &= \int_{t+u}^\infty e^{-\lambda s} E\{E\{f(x^{(a)}(s))/\boldsymbol{B}_{t+u}\}/\boldsymbol{B}_t\}ds\\
&= \int_{t+u}^\infty e^{-\lambda s} E\{f(x^{(a)}(s))/\boldsymbol{B}_t\}ds \quad (\because\ \boldsymbol{B}_{t+u} \supset \boldsymbol{B}_t)\\
&\leqslant \int_t^\infty e^{-\lambda s} E\{f(x^{(a)}(s))/\boldsymbol{B}_t\}ds = Y(t).
\end{aligned}
$$

因此, $Y(t)$ 对 $\boldsymbol{B}_t(t \geqslant 0)$ 是半鞅 (§34). 又

$$
\begin{aligned}
E\{Y(t)\} = e^{-\lambda t} E\{R_\lambda f(x^{(a)}(t))\} &= e^{-\lambda t} \int_R R_\lambda f(x) P(t, a, dx)\\
&= e^{-\lambda t} T_t R_\lambda f(a).
\end{aligned}
$$

由于 $x^{(a)}(t)$ 可分, 故 $R_\lambda f(x^{(a)}(t))$ 可分, 因此 $Y(t)$ 可分, 故利用 §34 定理 1 得

$$P(\mathfrak{A}(Y)) = 1,$$

这里

$$\mathfrak{A}(Y) = \text{“对一切 } t, Y(t-0) \text{ 和 } Y(t+0) \text{ 存在”}.$$

因此

$$P\{\mathfrak{A}(R_\lambda f(x^{(a)}(\cdot)))\} = 1.$$

由于 $\lambda R_\lambda f(x)$ 当 $\lambda \to \infty$ 时一致收敛于 $f(x)$, 故

$$P\{\mathfrak{A}(f(x^{(a)}(\cdot)))\} = 1. \tag{42.2}$$

设 $f_n(n = 1, 2, \cdots)$ 为 $\boldsymbol{C}$ 内非负函数列, 且分享 R 中不同的两点, 则映射

$$R \ni x \to \boldsymbol{f}(x) = (f_1(x), f_2(x), \cdots) \in \prod_n [0, \| f_n \|] \equiv K$$

是连续一对一的. 此处设 K 中的拓扑为弱拓扑. 由于 R 是紧致的 Hausdorff 空间, 故逆映射也连续, 这个对应为同胚. 因各个 f_n 都满足 (42.2), 又可数多个概率为 1 的集的交的概率也是 1, 故

$$P(\mathfrak{A}(f_n(x^{(a)}(\cdot))), n = 1, 2, \cdots) = 1.$$

由 K 的弱拓扑的定义, 有

$$P(\mathfrak{A}(\boldsymbol{f}(x^{(a)}(\cdot)))) = 1.$$

因为 $x \leftrightarrow \boldsymbol{f}(x)$ 是拓扑对应, 故

$$P(\mathfrak{A}(x^{(a)}(\cdot))) = 1.$$

定理的前半部分得到证明.

于是, 对 $f, g \in \boldsymbol{C}$, 由 (41.5) 得

$$E\{f(x^{(a)}(t))g(x^{(a)}(t+u))\} = E\{f(x^{(a)}(t))T_u g(x^{(a)}(t))\}.$$

如果令 $u \downarrow 0$, 则有

$$E\{f(x^{(a)}(t))g(x^{(a)}(t+0))\} = E\{f(x^{(a)}(t))g(x^{(a)}(t))\}.$$

因此, 对于 $R \times R$ 上的任意连续函数 h, 有

$$E\{h(x^{(a)}(t), x^{(a)}(t+0))\} = E\{h(x^{(a)}(t), x^{(a)}(t))\},$$

特别地, 如果令 h 为 R 上的有界距离函数, 则可得 (42.1).

对于可分 Markoff 过程 $x^{(a)}(t)$, 若定义

$$\widetilde{x}^{(a)}(t) = x^{(a)}(t+0),$$

则由上述定理知, $\widetilde{x}^{(a)}(t)$ 的样本过程至多只有第一类不连续点, 而且 $\widetilde{x}^{(a)}(t)$ 在弱的意义下 (§13) 与 $x^{(a)}(t)$ 一致. $f(t+0)$ 和 $f(t-0)$ 存在, 而 $f(t+0) = f(t) \neq f(t-0)$ 时, t 称为 f 的**第一类不连续点**. 因此, 不妨假定 Markoff 过程的样本过程至多只有第一类不连续点, 以后总设这假定成立. 又样本过程以概率 1 连续时, 则称为**扩散过程** (diffusion process).

§43 Markoff 过程 (3) 强 Markoff 性

在 §41 中, 已就 Markoff 性作了说明, 即若已知在某时刻 t 的状态, 则将来变动的概率法则与过去无关, 而好像从 t 时所处的状态出发一样. 这里 t 虽然是任意的, 但它却是一个常数, 下面将考虑 t 是随机变量的情形, 并研究同样的性质能否成立. 对于任意的随机变量 t 虽不成立, 但可以证明, 对于所谓 Markoff 时间的特殊的随机变量 $\tau(\omega)$ 是成立的.

定义 1　称 $\tau=\tau^{(a)}(\omega)$ 为 $x^{(a)}(t,\omega)(0\leqslant t<\infty)$ 的 **Markoff 时间** (Markoff time), 如果 τ 是取值于 $[0,\infty]$ 中的随机变量, 而且对于任意的 $t>0$,

$$\{\omega/\tau^{(a)}(\omega)<t\}\in \boldsymbol{B}_t(\text{即 } \boldsymbol{B}_t^{(a)}).$$

允许 $\tau=\infty$, 特别地, 如果 $P(\tau<\infty)=1$, 则称它为 **有限 Markoff 时间.**

前面对任意常数 t, 定义了 $\boldsymbol{B}_t$. 现在推广到 Markoff 时间 τ, 而有下面的定义.

定义 2　若 τ 是 $x^{(a)}(t)$ 的 Markoff 时间, 则以 $\boldsymbol{B}_\tau(=\boldsymbol{B}_\tau^{(a)})$ 表示包含所有形如 $E_\alpha\cap\{\omega/\tau\geqslant\alpha\},(\alpha>0,E_\alpha\in\boldsymbol{B}_\alpha)$ 的 ω 集的最小 Borel 体, 并称为 "由 $x^{(a)}(t)(t\leqslant\tau)$ 决定的 Borel 体".

(i) 强 Markoff 性的最简单的情形如下.

设 τ 为 $x^{(a)}(t)$ 的有限 Markoff 时间, 则

$$\begin{aligned}P(x^{(a)}(t+\tau)\in E/\boldsymbol{B}_\tau)&=P(x^{(b)}(t)\in E)_{b=x^{(a)}(\tau)}\\&=P(t,x^{(a)}(\tau),E),\end{aligned}$$

对有限 $(\boldsymbol{B}_R)$ 可测函数 $f(x)$, 有

$$E\{f(x^{(a)}(t+\tau))/\boldsymbol{B}_\tau\}=E\{f(x^{(b)}(t))\}_{b=x^{(a)}(\tau)}.$$

只需对 f 为连续函数的情形证明上式. 此时, 右边化为 $T_tf(x^{(a)}(\tau))$. 为了证明这一点, 只要证对于 $M\in\boldsymbol{B}_\tau$,

$$E\{f(x^{(a)}(\tau+u));M\}=E\{T_uf(x^{(a)}(\tau));M\}. \tag{43.1}$$

即可. 由于两边对 M 可加, 故又只需在 $M=E_\alpha\cap\{\omega/\alpha\leqslant\tau<\beta\}$ 时证明上式即可. 令

$$\alpha_{ni}=\alpha+\frac{i}{n}(\beta-\alpha),\quad \tau^{(n)}=\alpha_{ni}(\text{在 } \alpha_{n,i-1}\leqslant\tau<\alpha_{ni} \text{ 时}).$$

显然 $\tau^{(n)}\downarrow\tau$, 因而, 由于 $x^{(a)}(t)$ 的右连续性 (见上节) 得

$$\begin{aligned}x^{(a)}(\tau+u)&=\lim_{n\to\infty}x^{(a)}(\tau^{(n)}+u).\\A&=E\{f(x^{(a)}(\tau^{(n)}+u));E_\alpha\cap\{\alpha\leqslant\tau<\beta\}\}\\&=\sum_{i=1}^{n}E\{f(x^{(a)}(\alpha_{n,i}+u));E_\alpha\cap\{\alpha_{n,i-1}\leqslant\tau<\alpha_{n,i}\}\},\end{aligned}$$

由于 $\alpha \leqslant \alpha_{n,i-1} < \alpha_{n,i}$, 故 $E_\alpha, \{\omega/\tau < \alpha_{n,i-1}\}, \{\omega/\tau < \alpha_{n,i}\}$ 都属于 $\boldsymbol{B}_{\alpha_{n,i}}$, 所以 $E_\alpha \cap \{\alpha_{n,i-1} \leqslant \tau < \alpha_{n,i}\}$ 也属于 $\boldsymbol{B}_{\alpha_{n,i}}$. 因此

$$A = \sum_{i=1}^{n} E\{E\{f(x^{(a)}(\alpha_{n,i}+u))/\boldsymbol{B}_{\alpha_{n,i}}\}; E_\alpha \cap \{\alpha_{n,i-1} \leqslant \tau < \alpha_{n,i}\}\},$$

利用 Markoff 性

$$\begin{aligned} &= \sum_{i=1}^{n} E\{T_u f(x^{(a)}(\alpha_{n,i})); E_\alpha \cap \{\alpha_{n,i-1} \leqslant \tau < \alpha_{n,i}\}\} \\ &= E\{T_u f(x^{(a)}(\tau^{(n)})); E_\alpha \cap (\alpha \leqslant \tau < \beta)\}. \end{aligned}$$

因为 $T_t f(x)$ 与 $f(x)$ 都连续, 故令 $n \to \infty$ 即得 (43.1).

(ii) 推广上面的结果得

$$\begin{aligned} &P\{x^{(a)}(\tau+u_1) \in E_1, \cdots, x^{(a)}(\tau+u_n) \in E_n/\boldsymbol{B}_\tau\} \\ =&P\{x^{(b)}(u_1) \in E_1, \cdots, x^{(b)}(u_n) \in E_n\}_{b=x^{(a)}(\tau)}. \end{aligned}$$

更一般地, 对 $\Xi \in \boldsymbol{B}(R^{[0,\infty)})$,

$$P\{x^{(a)}(\tau+\cdot) \in \Xi/\boldsymbol{B}_\tau\} = P\{x^{(b)}(\cdot) \in \Xi\}_{b=x^{(a)}(\tau)}.$$

至于期望, 则对 $R^{[0,\infty)}$ 上的有界 ($\boldsymbol{B}(R^{[0,\infty)})$) 可测函数 f, 有

$$E\{f(x^{(a)}(\tau+\cdot))/\boldsymbol{B}_\tau\} = E\{f(x^{(b)}(\cdot))\}_{b=x^{(a)}(\tau)}.$$

只需证明后式. 若注意两边对于 f 是可加的, 则只要对 f 是一坐标上的连续函数的有限积的情况证明就够了. 这里只设 f 是两个函数的积, 一般情形, 可用同样方法证明.

要证的是: 对于 R 上的连续函数 f_1, f_2 及 $0 < u_1 < u_2$, 有

$$\begin{aligned} &E\{f_1(x^{(a)}(\tau+u_1))f_2(x^{(a)}(\tau+u_2))/\boldsymbol{B}_\tau\} \\ =&E\{f_1(x^{(b)}(u_1))f_2(x^{(b)}(u_2))\}_{b=x^{(a)}(\tau)}. \end{aligned}$$

证明方法同 (i), 但这时要注意的是

$$g_1(x) = T_{u_2-u_1} f_2(x), \quad g_2(x) = f_1(x) g_1(x), \quad g_3(x) = T_{u_1} g_2(x)$$

都是连续函数, 且若 $s_1 < s_2$, 则 $\boldsymbol{B}_{s_1} \subset \boldsymbol{B}_{s_2}$, 因此

$$E(E(\cdot/\boldsymbol{B}_{s_2})/\boldsymbol{B}_{s_1}) = E(\cdot/\boldsymbol{B}_{s_1}).$$

(iii) 利用强 Markoff 性, 可得到下述常用的性质. 设 $\tau=\tau^{(a)}(\omega)$ 是 $x^{(a)}(t)$ 的有限 Markoff 时间, 令

$$\begin{aligned}
P^0(t,a,\mathrm{d}y) &= P(x^{(a)}(t)\in \mathrm{d}y, \tau>t),\\
T_t^0 f(a) &= \int_R f(y)P^0(t,a,\mathrm{d}y),\\
R_\lambda^0 f(a) &= \int_0^\infty \mathrm{e}^{-\lambda t}T_t^0 f(a)\mathrm{d}t,\\
\varphi^{(a)}(\mathrm{d}s,\mathrm{d}y) &= P(\tau\in \mathrm{d}s, x^{(a)}(\tau)\in \mathrm{d}y),\\
\widehat{\varphi}^{(a)}(\lambda,\mathrm{d}y) &= \int_0^\infty \mathrm{e}^{-\lambda s}\varphi(\mathrm{d}s,\mathrm{d}y),
\end{aligned}$$

则

$$R_\lambda f(a)=R_\lambda^0 f(a)+\int_{y\in R} R_\lambda f(y)\widehat{\varphi}^{(a)}(\lambda,\mathrm{d}y), \tag{43.2}$$

$$T_t f(a)=T_t^0 f(a)+\int_{y\in R}\int_{s=0}^t T_{t-s}f(y)\varphi^{(a)}(\mathrm{d}s,\mathrm{d}y). \tag{43.3}$$

设 $[0,s)$ 的示性函数为 $c_s(t)$, 则

$$\begin{aligned}
E\{f(x^{(a)}(t));\tau>t\} &= E\{f(x^{(a)}(t))c_\tau(t)\}\\
&= \int_R f(y)P^0(t,a,\mathrm{d}y) = T_t^0 f(a),\\
E\left\{\int_0^\tau \mathrm{e}^{-\lambda t}f(x^{(a)}(t))\mathrm{d}t\right\} &= E\left\{\int_0^\infty \mathrm{e}^{-\lambda t}f(x^{(a)}(t))c_\tau(t)\mathrm{d}t\right\}\\
&= \int_0^\infty \mathrm{e}^{-\lambda t}E\{f(x^{(a)}(t))c_\tau(t)\}\mathrm{d}t\\
&= \int_0^\infty \mathrm{e}^{-\lambda t}T_t^0 f(a)\mathrm{d}t = R_\lambda^0 f(a).\\
R_\lambda f(a) &= \int_0^\infty \mathrm{e}^{-\lambda t}E\{f(x^{(a)}(t))\}\mathrm{d}t\\
&= E\left\{\int_0^\infty \mathrm{e}^{-\lambda t}f(x^{(a)}(t))\mathrm{d}t\right\}\\
&= E\left\{\int_0^\tau \mathrm{e}^{-\lambda t}f(x^{(a)}(t))\mathrm{d}t\right\}\\
&\quad + E\left\{\int_\tau^\infty \mathrm{e}^{-\lambda t}f(x^{(a)}(t))\mathrm{d}t\right\}.
\end{aligned}$$

$$第 1 项 = R_\lambda^0 f(a).$$

$$\begin{aligned}
第 2 项 &= E\left\{\mathrm{e}^{-\lambda\tau}\int_0^\infty \mathrm{e}^{-\lambda t} f(x^{(a)}(\tau+t))\mathrm{d}t\right\}\\
&= E\left\{\mathrm{e}^{-\lambda\tau}E\left\{\int_0^\infty \mathrm{e}^{-\lambda t} f(x^{(a)}(\tau+t))\mathrm{d}t/\boldsymbol{B}_\tau\right\}\right\}\\
&= E\left\{\mathrm{e}^{-\lambda\tau}\int_0^\infty \mathrm{e}^{-\lambda t} E\{f(x^{(a)}(\tau+t))/\boldsymbol{B}_t\}\mathrm{d}t\right\}\\
&= E\left\{\mathrm{e}^{-\lambda\tau}\int_0^\infty \mathrm{e}^{-\lambda t} T_t f(x^{(a)}(\tau))\mathrm{d}t\right\}\\
&= E\{\mathrm{e}^{-\lambda\tau}R_\lambda f(x^{(a)}(\tau))\}\\
&= \int_{t=0}^\infty\int_{y\in R} \mathrm{e}^{-\lambda t}R_\lambda f(y)\varphi^{(a)}(\mathrm{d}t,\mathrm{d}y)\\
&= \int_{y\in R} R_\lambda f(y)\widehat{\varphi}^{(a)}(\lambda,\mathrm{d}y),
\end{aligned}$$

故 (43.2) 得证. 若取 (43.3) 右边 (关于 t) 的 Laplace 变换, 则得 (43.2) 的右边, 因而等于 $R_\lambda f(a)$. 由于 $R_\lambda f(a)$ 等于 $T_t f(a)$ 的 Laplace 变换, 故 (43.3) 得证.

§44 Markoff 时间

列举 Markoff 时间的性质以及例子如下.

(i) 若 τ_1,τ_2 是 Markoff 时间, 则 $\max(\tau_1,\tau_2)$, $\min(\tau_1,\tau_2)$ 也是 Markoff 时间. 因为

$$\{\omega/\max(\tau_1,\tau_2)<t\}=\{\omega/\tau_1<t\}\cap\{\omega/\tau_2<t\},$$

$$\{\omega/\min(\tau_1,\tau_2)<t\}=\{\omega/\tau_1<t\}\cap\{\omega/\tau_2<t\}.$$

(ii) 若 $\tau_1\leqslant\tau_2\leqslant\cdots$ 都是 Markoff 时间, 则 $\tau=\lim\tau_n$ 也是 Markoff 时间. 因为

$$\{\omega/\tau<t\}=\bigcup_n\bigcap_m\left\{\omega/\tau_m<t-\frac{1}{n}\right\}.$$

(iii) 若 $\tau_1\geqslant\tau_2\geqslant\cdots$ 都是 Markoff 时间, 则 $\tau=\lim\ \tau_n$ 也是 Markoff 时间. 因为

$$\{\omega/\tau<t\}=\bigcup_n\{\omega/\tau_n<t\}.$$

(iv) $\tau(\omega)\equiv t$ 是 Markoff 时间.

(v) **Dynkin 引理** 若 τ 是有限 Markoff 时间, 则

$$f(a)=-E\int_0^{\tau}Af(x^{(a)}(t))\mathrm{d}t+Ef(x^{(a)}(\tau)),\quad f\in\mathfrak{D}(A).$$

证明 令 $h_\lambda=\lambda f-Af$, 则 $f=R_\lambda h_\lambda$.

$$\begin{aligned}f(a)=R_\lambda h_\lambda(a)&=\int_0^{\infty}\mathrm{e}^{-\lambda t}T_t h_\lambda(a)\mathrm{d}t\\&=E\left\{\int_0^{\infty}\mathrm{e}^{-\lambda t}h_\lambda(x^{(a)}(t))\mathrm{d}t\right\}\\&=E\left\{\int_0^{\tau}\right\}+E\left\{\int_{\tau}^{\infty}\right\}=A_\lambda+B_\lambda.\end{aligned}$$

因为当 $\lambda\to 0$ 时, $h_\lambda\to -Af$, 所以

$$A_\lambda\to -E\int_0^{\tau}Af(x^{(a)}(t))\mathrm{d}t,$$

$$\begin{aligned}B_\lambda&=E\left\{\mathrm{e}^{-\lambda\tau}\int_0^{\infty}\mathrm{e}^{-\lambda t}h_\lambda(x^{(a)}(\tau+t))\mathrm{d}t\right\}\\&=E\left\{\mathrm{e}^{-\lambda\tau}\int_0^{\infty}\mathrm{e}^{-\lambda t}E(h_\lambda(x^{(a)}(\tau+t))/\boldsymbol{B}_\tau)\mathrm{d}t\right\}\\&=E\left\{\mathrm{e}^{-\lambda\tau}\int_0^{\infty}\mathrm{e}^{-\lambda t}T_t h_\lambda(x^{(a)}(\tau))\mathrm{d}t\right\}\\&=E\{\mathrm{e}^{-\lambda\tau}R_\lambda h_\lambda(x^{(a)}(\tau))\}\\&=E\{\mathrm{e}^{-\lambda\tau}f(x^{(a)}(\tau))\}\to E\{f(x^{(a)}(\tau))\}.\end{aligned}$$

(vi) 令 F 为 R 的闭集. $x^{(a)}(t)$ 跑出 F 的时刻的下限称为 F 的**最初通过时间** (first passage time), 记为 $\tau_F=\tau_F^{(a)}(\omega)\cdot x^{(a)}(t)$ 在 τ_F 以前处于 F 上, 在 τ_F 与 $\tau_F+\varepsilon$ 间一定会自 F 跑出. 在 τ_F 时 $x^{(a)}(t)$ 是否位于 F 上尚不知道. 若 $x^{(a)}(t)$ 在 $t=\tau_F$ 时连续, 则 $x^{(a)}(\tau_F)$ 处于 F 的边界 (包含于 F) 上, τ_F 是 Markoff 时间. 因为

$$\begin{aligned}\{\omega/\tau_F<t\}&=\bigcup_{s<t}\{\omega/x^{(a)}(s,\omega)\in F^{\mathrm{c}}\}\\&=\bigcup_{r<t}\{\omega/x^{(a)}(r,\omega)\in F^{\mathrm{c}}\}\end{aligned}$$

(这里 r 是有理数). 由于 F^c 是开集, 且 $x^{(a)}(s,\omega)$ 右连续, 故后一等号显然成立. 右边项是属于 $\boldsymbol{B}_t$ 的集的可数和, 故仍属于 $\boldsymbol{B}_t$.

(vii) 令 U 为 R 的开集. U 的最初通过时间 τ_U 可与 τ_F 同样定义. 特别地, 令 $x^{(a)}(t)$ 为 R 上的扩散过程.

根据 R 的拓扑性质, 考虑由内部逼近 U 的闭集列

$$F_1 \subset F_2 \subset \cdots \subset F_n \subset \cdots \to U$$

便得

$$\tau_{F_1} \leqslant \tau_{F_2} \leqslant \cdots \to \tau_U.$$

故由 (vi) 与 (ii) 可知, τ_U 也是 Markoff 时间.

(viii) 在上述 (vi) 中, 特别当 F 是一点 a 时, 试考虑 $\tau = \tau_F^{(a)}(\omega)$. 此为从 a 出发的 $x^{(a)}(t)$ 最初离开 a 的时间, 或者也可叫做停留于 a 的时间. 试求 τ 的分布. 设

$$p(t) = P(\tau > t),$$

利用 Markoff 性可得

$$p(t+s) = p(t)p(s).$$

实际上,

$$\begin{aligned}
p(t+s) &= P(x^{(a)}(u) = a, 0 \leqslant u \leqslant t+s) \\
&= P(x^{(a)}(u) = a, 0 \leqslant u \leqslant t; x^{(a)}(t+v) = a, 0 \leqslant v \leqslant s) \\
&= E\{P(x^{(a)}(t+v) = a, 0 \leqslant v \leqslant s/\boldsymbol{B}_t); x^{(a)}(u) = a, 0 \leqslant u \leqslant t\} \\
&= E\{P(x^{(b)}(v) = a, 0 \leqslant v \leqslant s)_{b=x^{(a)}(t)}; x^{(a)}(u) = a, 0 \leqslant u \leqslant t\} \\
&= E\{P(x^{(a)}(v) = a, 0 \leqslant v \leqslant s); x^{(a)}(u) = a, 0 \leqslant u \leqslant t\} \\
&= p(s) \cdot p(t).
\end{aligned}$$

因此, 除去 $p(t) \equiv 0$ 的情形, 常有 $p(t) > 0, p(t) \leqslant 1$, 在后一情形, $p(t) = \mathrm{e}^{-\lambda t}(\lambda \geqslant 0)$. 当 $\lambda = 0$ 时, $p(t) \equiv 1$. 当 $\lambda > 0$ 时, 当 t 由 0 增至 ∞ 时 $p(t)$ 由 1 减至 0. 当 $p(t) \equiv 0$ 时, 从 a 出发, 在下一瞬时即自 a 跑出. 亦即, 无论 ε 为如何小的正数, 在 $0 < t < \varepsilon$ 之间, 总会有些时刻离开 a. 此时 a 称为**瞬时状态** (instantaneous state), 或称为**瞬时逗留点**.

当 $p(t) \equiv 1$ 时, 则从 a 出发, 永远不能离开 a. 此时称 a 为**套点** (trap).

当 $p(t) = \mathrm{e}^{-\lambda t}(\lambda > 0)$ 时, τ 的分布是**指数型** (exponential type), τ 称

为指数型逗留时间 (exponential holding time, exponential waiting time), a 称为**指数型逗留点.**

下列四条件相互等价.

(A) a 是套点;　　(B) $P(t,x,E)=\delta(a,E), t>0$;

(C) $T_t f(a)=f(a), t>0, f\in \boldsymbol{C}$; (D) $Af(a)=0, f\in\mathfrak{D}(A)$.

(A)→ (B) → (C) → (D) 显然. 若假定 (D), 则对于 $f\in\mathfrak{D}(A)$, 因有 $T_t f\in\mathfrak{D}(A)$, 故

$$\frac{\mathrm{d}T_t f(a)}{\mathrm{d}t}=AT_t f(a)=0,$$

且得 $T_t f(a)=f(a)$. 因 $\overline{\mathfrak{D}(A)}=\boldsymbol{C}$, 故得 (C). (C)→(B) 显然. 若假定 (B), 则

$$P(x^{(a)}(t)=a)=1,\quad t>0$$

故 $P(x^{(a)}(t)=a, t$ 为正有理数$)$=1. 因为 $x^{(a)}(t)$ 右连续, 故 $P(x^{(a)}(t)=a, t>0)=1$, 这表示 a 为套点.

(ix) 扩散过程没有指数型逗留点.

证明　若 a 不是套点, 则有 $f\in\mathfrak{D}(A)$ 使 $Af(a)\neq 0$. 设离开 a 的最初时间为 τ, 则 τ 是有限 Markoff 时间. 因扩散过程是连续的, 所以 $x^{(a)}(\tau)=a$. 故由 (v) 中 Dynkin 引理推出

$$f(a)=-Af(a)\cdot E(\tau)+f(a),\quad \text{即}\quad E(\tau)=0,$$

因此 $P(\tau=0)=1$. 这表示 a 是瞬时逗留点.

(x) 若 a 是 M(闭或开集) 的内点, 则 $E(\tau_M^{(a)})>0$.

证明　若 $E(\tau_M^{(a)})=0$, 则 $P(\tau_M^{(a)}=0)$=1. 因此

$$P(x^{(a)}(0+)\notin M^i)=1.$$

然而 $P(x^{(a)}(0+)=a)=1$, 与 $a\in M^i$ 矛盾.

(xi) 若 a 不是套点, 则有 a 的邻域 U, 使对于 $b\in U, E(\tau_U^{(b)})$ 为一致有界. 特别在扩散过程时, 对任意 ε, 可适当地选取 U, 使 $E(\tau_U^{(b)})<\varepsilon, b\in U$. 但后一性质对一般的 Markoff 过程未必成立.

证明　因 a 不是套点, 故存在 $f\in\mathfrak{D}(A)$ 使 $Af(a)\neq 0$. 如有必要可取 $-f$ 而得 $Af(a)>0$, 因 Af 连续, 故存在正数 α 及 a 的领域 U, 使 $Af(b)>\alpha, b\in U$. 由上述 Dynkin 引理得

$$f(b)\leqslant -\alpha E(\tau_U^{(b)})+\|f\|,\quad \text{即}\quad E(\tau_U^{(b)})\leqslant\frac{1}{\alpha}(\|f\|-f(b))<\infty.$$

特别对扩散过程, 有 $x^{(b)}(\tau_U^{(b)}) \in \overline{U}$. 由于 f 的连续性, f 在 $\overline{U}$ 中的变差可以限制得小于 $\alpha\varepsilon$. 故 $f(x^{(b)}(\tau_U^{(b)})) < f(b) + \alpha\varepsilon$. 由 Dynkin 引理得

$$f(b) < -\alpha E(\tau_U^{(b)}) + (f(b) + \alpha\varepsilon), \quad 故 \quad E(\tau_U^{(b)}) < \varepsilon.$$

对一般情形未必如此, 例如考虑指数型逗留点 a, 无论 U 取得多么小, $\tau_U^{(a)}$ 总比 a 的逗留时间 τ 大, 且因 $E(\tau) > 0$, 故不能使 $E(\tau_U^{(a)})$ 小于 $E(\tau)$. 含有指数型逗留点的 Markoff 过程的存在性见后.

§45 Dynkin 关于生成算子的定理

前已说明: "转移概率 $P(t, x, E) \leftrightarrow$ 半群 $T_t \leftrightarrow$ Markoff 过程 $x^{(a)}(t)$", 这个对应关系是一对一的. T_t 的生成算子 A 既叫做 $P(t, x, E)$ 的生成算子, 又叫做 $x^{(a)}(t)$ 的生成算子. 利用 $x^{(a)}(t)$ 来定义 A 就是下述的 Dynkin 定理.

定理 1 令 U 为 α 的领域, $\tau = \tau_U^{(a)}$. 设

$$Df(U) = \frac{E(f(x^{(a)}(\tau))) - f(a)}{E(\tau)},$$

(当 $E(\tau) = \infty$ 时, 令 $Df(U) = 0$). 则对 $f \in \mathfrak{D}(A)$ 有

$$Df(U) \to Af(a) \quad (U \to a). \tag{45.1}$$

又若当 $U \to a$ 时, $Df(U)$ 逼近于 a 的连续函数, 则 $f \in \mathfrak{D}(A)$, 从而上面的 (45.1) 成立.

$Df(U) \to Af(a)$ 的详细内容是: 如取 U 足够小, 则

$$|Df(U) - Af(a)| \leqslant \sup_{b \in \overline{U}} |Af(b) - Af(a)|. \tag{45.2}$$

证明 若 a 是套点, 则 (45.2) 的左边为 0, 故成立. 若 a 不是套点, 则由于 τ 是有限 Markoff 时间, 故由 Dynkin 引理, 有

$$f(a) = -E\int_0^\tau Af(x^{(a)}(t))\mathrm{d}t + Ef(x^{(a)}(\tau)).$$

因 a 不是套点, 故当 U 充分小时, 有 $E(\tau) < \infty$, 又因 a 是 U 的内点, 故

$E(\tau) > 0$. 由上式得

$$\begin{aligned}
&\left|\frac{Ef(x^{(a)}(\tau)) - f(a)}{E(\tau)} - Af(a)\right| \\
\leqslant &\frac{1}{E(\tau)} E\int_0^\tau |Af(x^{(a)}(t)) - Af(a)|\mathrm{d}t \\
\leqslant &\sup_{b\in\overline{U}} |Af(b) - Af(a)| \to 0 \qquad (U \to a).
\end{aligned}$$

以 $\widetilde{\mathfrak{D}}$ 表示使 $Df(U)$ 逼近于 a 的连续函数 f 的全体, 对这样的 f 定义 $\lim Df(U)$ 为 $\widetilde{A}f$, 由上述得 $\widetilde{A} \supset A$. 现在指出 $\lambda - \widetilde{A}$ 是一一对应的. 为此只需由 $\widetilde{A}u = \lambda u$ 引出 $u = 0$. 若令 $u(x)$ 的最小值为 $u(a)$, 则由 $\widetilde{A}$ 的定义得 $Au(a) \geqslant 0$, 故 $\lambda u(a) \geqslant 0$, 即 $u(a) \geqslant 0$, 且恒有 $u(x) \geqslant 0$. 又因 $\widetilde{A}(-u) = \lambda(-u)$, 故 $-u(x) \geqslant 0$, 于是 $u(x) \equiv 0$. 从而有 $(\lambda - \widetilde{A})^{-1}$ 存在. 又由 $\widetilde{A} \supset A$ 得

$$(\lambda - \widetilde{A})^{-1} \supset (\lambda - A)^{-1} = R_\lambda, \quad \mathfrak{D}(R_\lambda) = \boldsymbol{C},$$

故 $\supset$ 变为等号 $=$, 因此 $\widetilde{A} = A$, 故得 $\widetilde{\mathfrak{D}} = \mathfrak{D}(A)$.

今将上述的 $Df(U)$ 变成其他形式, 设

$$\pi_U(a, \mathrm{d}y) = P(x^{(a)}(\tau) \in \mathrm{d}y).$$

这测度的支撑点显然包含于 U^c(尤其是扩散过程时, 则包含于 U 的边界 ∂U). 有

$$Df(U) = \left\{\int_{U^c} \pi_U(a, \mathrm{d}y) f(y) - f(a)\right\} / E(\tau). \tag{45.3}$$

又为了改写 $E(\tau)$, 引进记号 $p_U(a) = E(\tau_U^{(a)})$($a \in U^c$ 时, $\tau_U^{(a)} \equiv 0$, 因此令 $p_U(a) = 0$).

令 $U \subset V$. 试证

$$p_V(a) = p_U(a) + \int \pi_U(a, \mathrm{d}y) p_V(y). \tag{45.4}$$

在 $\mathbb{R}^{[0,\infty)}$ 上能取 $(\boldsymbol{B}(\mathbb{R}^{[0,\infty)}))$ 可测函数 ϕ_V, 使

$$\tau_V^{(a)} = \phi_V(x^{(a)}(\cdot)).$$

为此, 首先对闭集 F 令

$$\phi_F(\xi) = \inf\{t/\xi(t) \in F^c, t \text{ 为有理数}\},$$

则由 $x^{(a)}(t)$ 的右连续性与 F^c 是开集, 得 $\tau_F^{(a)} = \phi_F(x^{(a)}(\cdot))$. 对开集 V, 取由内部逼近于此开集的闭集列 $F_1 \subset F_2 \subset \cdots \to V$, 并令

$$\phi_V(\xi) = \lim_{n\to\infty} \phi_{F_n}(\xi),$$

则得 $\tau_V^{(a)} = \phi_V(x^{(a)}(\cdot))$, 且

$$\tau_V^{(a)} = \tau_U^{(a)} + \tau_V^{(a)} - \tau_U^{(a)} = \tau_U^{(a)} + \phi_V(x^{(a)}(\tau_U^{(a)} + \cdot)).$$

考虑强 Markoff 性得

$$\begin{aligned} E(\tau_V^{(a)}) &= E(\tau_U^{(a)}) + E\{\phi_V(x^{(a)}(\tau_U^{(a)} + \cdot))\} \\ &= E(\tau_U^{(a)}) + E\{E\{\phi_V(x^{(a)}(\tau_U^{(a)} + \cdot))/\boldsymbol{B}\tau_U^{(a)}\}\} \\ &= E(\tau_U^{(a)}) + E\{E\{\phi_V(x^{(b)}(\cdot))\}_{b=x^{(a)}(\tau_U^{(a)})}\} \\ &= E(\tau_U^{(a)}) + E\{E(\tau_V^{(b)})_{b=x^{(a)}(\tau_U^{(a)})}\}, \end{aligned}$$

即

$$\begin{aligned} p_V(a) &= p_U(a) + E\{p_V(x^{(a)}(\tau_U^{(a)}))\} \\ &= p_U(a) + \int p_V(y)\pi_U(a, \mathrm{d}y). \end{aligned}$$

由此得证 (45.4). 因为上述的 $E(\tau)$ 即 $p_U(a)$, 故

$$E(\tau) = p_V(a) - \int \pi_U(a, \mathrm{d}y)p_V(y).$$

由此得

$$Df(U) = \frac{\int \pi_U(a, \mathrm{d}y)f(y) - f(a)}{\int \pi_U(a, \mathrm{d}y)(-p_V(y)) - (-p_V(a))}. \tag{45.5}$$

§46 Markoff 过程的例子

下面将应用以上数节的理论, 来重新讨论 §35 及 §40 所列举的对应于转移概率的 Markoff 过程的例子.

例 1 有限个状态的 Markoff 过程 (Markoff process with finite states) 首先考虑 §35 的例 1(或 §40 的例 1). 记号如前. 但用 $\{1, 2, \cdots, n\}$ 表示 R 中的点. 因为 $x^{(i)}(t)$ 的样本过程是从 i 出发, 且是至多只有第一

类不连续点的函数, 又所取的值只是 $1,2,\cdots,n$, 故 $x^{(i)}(t)$ 必为阶梯函数. 若自 i 出发, 作如下运动: 在长为 τ 的一段时间内, 停留于 i, 然后转移到另一 j, 再在此停留一段时间, 又转移到另一 k. 因只由一点 i 所构成的集可以看作开集, 故得 $E(\tau)>0$, 即 i 不能为瞬时逗留点. 故 i 或为指数型逗留点, 或为套点. 若为套点, 则不论 f 如何, $Af(i)$ 总为 0, 故得

$$a_{ij}=0,\quad j=1,2,\cdots,n.$$

若为指数型逗留点, 则 $P(\tau>t)=\mathrm{e}^{-\lambda_i t}, E(\tau)=1/\lambda_i$. 由此可证 λ_i 和 a_{ij} 的关系为

$$\lambda_i=\sum_{j\neq i}a_{ij}=-a_{ii}. \tag{46.1}$$

为此, 令 t 充分小, 由 A 的定义有

$$p(t,i,i)=1+a_{ii}t+o(t). \tag{46.2}$$

又由 τ 的定义得

$$p(t,i,i)=\mathrm{e}^{-\lambda_i t}+\varepsilon. \tag{46.3}$$

此处 ε 是曾离开 i 再回到原处, 且在 t 时处于 i 的概率, 此事件之出现至少需要二次跳跃, 故其数量级至少为 $O(t^2)$(用强 Markoff 性可得到严格证明). 故 (46.3) 变为

$$p(t,i,i)=1-\lambda_i t+o(t). \tag{46.4}$$

以此与 (46.2) 比较即得 $\lambda_i=-a_{ii}$. 由 $x^{(i)}(t)$ 的右连续性, 可知 $x^{(i)}(\tau)$ 是从 i 经第一次转移后所处的新状态, 故其值 j 不同于 i, 其概率为

$$P(x^{(i)}(\tau)=j)=a_{ij}/\lambda_i. \tag{46.5}$$

为此, 考虑

$$P\{x^{(i)}(\tau)=j,\ \tau<T\}=\sum_{k=1}^{n}P\left\{x^{(i)}(\tau)=\frac{k-1}{n}T\leqslant\tau<\frac{k}{n}T\right\}.$$

对足够大的 n, 在 $\left[\frac{k-1}{n}T,\frac{k}{n}T\right]$ $(k=1,2,\cdots,n)$ 中的任一个区间内发生不少于二次跳跃的概率非常小 (因为 $x^{(i)}(t)$ 是阶梯函数, 故在长为 T 的时间内的跳跃次数是有限的), 故

$$=\sum_{k=1}^{n}P\left\{x^{(i)}(t)\equiv i,\ t\leqslant\frac{k-1}{n}T;\ x^{(i)}\left(\frac{k}{n}T\right)=j\right\}+o(1).$$

由 Markoff 性

$$=\sum_{k=1}^{n}P\Big\{x^{(i)}(t),\ t\leqslant\frac{k-1}{n}T\Big\}P\Big\{x^{(i)}\Big(\frac{T}{n}\Big)=j\Big\}+o(1).$$

但

$$P\Big(x^{(i)}\Big(\frac{T}{n}\Big)=j\Big)=p\Big(\frac{T}{n},i,j\Big)=a_{ij}\frac{T}{n}+o\Big(\frac{1}{n}\Big),$$

故

$$\begin{aligned}&=\sum_{k=1}^{n}\mathrm{e}^{-\lambda_i\frac{k-1}{n}T}a_{ij}\frac{T}{n}+no\Big(\frac{1}{n}\Big)+o(1)\\&\to\int_0^T\mathrm{e}^{-\lambda_i t}a_{ij}\mathrm{d}t=\frac{a_{ij}}{\lambda_i}(1-\mathrm{e}^{-\lambda_i T}).\end{aligned}$$

令 $T\to\infty$ 即得 (46.5).

于是, $x^{(i)}(t)$ 的运动如下进行. 自 i 出发, 经过指数型逗留时间 (其平均值为 $\lambda_i^{-1}=(\sum\limits_{j\neq i}a_{ij})^{-1}$) 后, 以概率 a_{ij}/λ_i 转入 j. 如果一切 $a_{ij}=0(j\neq i)$, 则 i 是套点. 否则, 转入 j 后又同样地运动以转入另一 k. 如一旦转入套点, 则永久停留于其上.

利用 Dynkin 定理, 可立刻求出 (46.1) 与 (46.5). i 点本身可以视为逼近于 i 的开集. 只需考虑 i 不是套点的情况. 因已知在 i 的逗留时间 τ 是指数型, 故令其平均值为 λ_i^{-1} 后, 其分布为 $P(\tau>t)=\mathrm{e}^{-\lambda_i t}$. 若令 $P(x^{(i)}(\tau)=j)$ 为 π_{ij}, 则这相当于 Dynkin 的 $\pi_U(a,\mathrm{d}y)$. 故

$$Af(i)=\frac{\sum\limits_{j\neq i}\pi_{ij}f(j)-f(i)}{\lambda_i^{-1}}=\sum_{j\neq i}\lambda_i\pi_{ij}f(j)-\lambda_i f(i).$$

另一方面

$$Af(i)=\sum_{j\neq i}a_{ij}f(j)-a_{ii}f(i).$$

二式比较即得

$$\lambda_i=-a_{ii}\Big(=\sum_{j\neq i}a_{ij}\Big),\quad \pi_{ij}=\frac{a_{ij}}{\lambda_i}.$$

此即 (46.1) 和 (46.5).

例 2 Wiener 过程 令 $B(t,\omega)$ 为 Wiener 过程, 并令 $B(0,\omega)\equiv 0$. 若定义

$$x^{(a)}(t,\omega)=a+B(t,\omega),$$
$$x^{(\infty)}(t,\omega)\equiv\infty,$$

则得 $R=\mathbb{R}^1\vee\{\infty\}$ 上一族随机过程, 它是以

$$P(t,x,E)=\int_E N_t(y-x)\mathrm{d}y$$

为转移概率的 Markoff 过程. 以后称此过程为由 Wiener 过程导出的 Markoff 过程, 或简称为 Wiener 过程, 它是扩散过程.

例 3 以 0 为反射壁 (reflecting barrier) 的 Wiener 过程 对前例中的 $x^{(a)}(t)$ 定义

$$y^{(a)}(t)=|x^{(a)}(t)|,\quad a\in[0,\infty].$$

此为与 §35 例 3(§40 例 3) 对应的 Markoff 过程, 也是扩散过程.

例 4 圆周上的旋转 考虑在周长为 1 的圆周上以速度 1 向正向旋转的运动. 圆周上的点由 mod 1 的实数决定, 而此运动则由

$$x^{(a)}(t,\omega)=a+t(\mathrm{mod}\ 1)$$

决定, 它是与

$$P(t,a,E)=\delta(a+t,E)$$

对应的 Markoff 过程 (见 §40 例 4). 这也是扩散过程. 若给定的初始值为 a, 则 $x^{(a)}(t,\omega)$ 等于 $a+t(\mathrm{mod}\ 1)$ 而与 ω 无关, 此即所谓决定性的 (deterministic) 过程.

例 5 若以 mod 1 来考虑 Wiener 过程 (例 2), 则得圆周上的 Markoff 过程, 称它为圆周上的 Wiener 过程. 这也是扩散过程 (见 §40 例 5).

§47 对时间为齐次的可加过程

在上节例 2 中已说过, Wiener 过程可看成 $R=\mathbb{R}^1\vee\{\infty\}$ 上的 Markoff 过程, 这可推广到一般的对时间为齐次的可加过程.

今设 $y(t,\omega)(0 \leqslant t \leqslant \infty)$ 为对时间为齐次的可加过程, $y(0,\omega)=0$. 令 $y(t,\omega)$ 的分布为 Φ_t. 对任意 $u>0$, 这个分布就是 $y(u+t,\omega)-y(u,\omega)$ 的分布. 由可加过程的性质有

$$\Phi_{t+s} = \Phi_t * \Phi_s. \tag{47.1}$$

Φ_t 为无穷可分分布, 其特征函数 $\varphi_t(z)$ 由

$$\left.\begin{aligned} &\varphi_t(z) = \mathrm{e}^{t\psi(z)}, \\ &\psi(z) = imz - \frac{v}{2}z^2 + \int_{-\infty}^{+\infty}\left(\mathrm{e}^{izu} - 1 - \frac{izu}{1+u^2}\right)n(\mathrm{d}u) \end{aligned}\right\} \tag{47.2}$$

决定, 此处 m 是实数, $u \geqslant 0$, n 是满足

$$\int_{-\infty}^{+\infty} \frac{u^2}{1+u^2} n(\mathrm{d}u) < \infty \tag{47.3}$$

的测度.

若对 $a \in R = \mathbb{R}^1 \vee \{\infty\}$ 定义

$$\begin{aligned} x^{(a)}(t,\omega) &= a + y(t,\omega), \\ x^{(\infty)}(t,\omega) &= \infty, \end{aligned}$$

则得 Markoff 过程, 其转移概率为

$$P(t,a,E) = \Phi_t(E(-)a), \quad E(-)a = \{\xi - a/\xi \in E\}. \tag{47.4}$$

若对此转移概率写出 Chapman-Kolmogoroff 方程, 则得 (47.1).

此 Markoff 过程的半群为

$$T_t f(x) = \int \Phi_t(\mathrm{d}y(-)x) f(y) = \int \Phi_t(\mathrm{d}y) f(y+x), \tag{47.5}$$

Fourier 变换的算子 $\mathfrak{F}$ 定义为

$$\mathfrak{F}g(z) = \lim_{n\to\infty} \int_{-n}^{n} \mathrm{e}^{izx} g(x)\mathrm{d}x, \quad \mathfrak{F}\mathrm{d}\mu(z) = \lim_{n\to\infty}\int_{-n}^{n} \mathrm{e}^{izx}\mathrm{d}\mu(x).$$

在广义函数的意义下取极限, $\mathfrak{F}$ 是广义函数意义下的 Fourier 变换. 对 (47.5) 两边施行 $\mathfrak{F}$ 得

$$\mathfrak{F}T_t f(x) = \varphi_t(-z)\mathfrak{F}f(z) = \mathrm{e}^{t\psi(-z)}\mathfrak{F}f(z). \tag{47.6}$$

施行 Fourier 变换后, 可见此半群有极简单的形状. 其次, 对上式两边就 t 施行 Laplace 变换. 因

$$\Re\psi(-z) = -\frac{v}{2}z^2 + \int_{-\infty}^{+\infty} (\cos zu - 1)n(\mathrm{d}u) \leqslant 0,$$

故右边可以简单地施行 Laplace 变换, 左边则因 $\mathfrak{F}$ 可与 t 的 Laplace 变换先后交换, 故得

$$\mathfrak{F}R_\lambda f(z) = \frac{1}{\lambda - \psi(-z)}\mathfrak{F}f(z). \tag{47.7}$$

又由 (47.6) 得

$$\mathfrak{F}Af(z) = \psi(-z)\mathfrak{F}f(z),$$

即

$$\mathfrak{F}Af(z) = \left\{ -imz - \frac{v}{2}z^2 + \int \left(\mathrm{e}^{-izu} - 1 + \frac{izu}{1+u^2}\right)n(\mathrm{d}u)\right\}\mathfrak{F}f(z). \tag{47.8}$$

由 Fourier 变换的性质得

$$\mathfrak{F}f'(z) = -iz\mathfrak{F}f(z), \quad \mathfrak{F}f''(z) = -z^2\mathfrak{F}f(z),$$
$$\mathfrak{F}\Delta_u f(z) = (\mathrm{e}^{-izu} - 1)\mathfrak{F}f(z), \quad \text{此处 } \Delta_u f(x) = f(x+u) - f(x),$$

因此, 若形式地求 (47.8) 两边的 Fourier 逆变换, 则得

$$Af(x) = mf'(x) - \frac{v}{2}f''(x) + \int \left(f(x+u) - f(x) - \frac{u}{1+u^2}f'(x)\right)n(\mathrm{d}u). \tag{47.9}$$

欲求 $\mathfrak{D}(A)$ 以明确地规定 A, 则相当麻烦.

§48 生灭过程

设有一团细菌, 其个数按下述法则变化: 设一个细菌在 $\mathrm{d}t$ 时间内分裂为 2 个的概率为 $p\mathrm{d}t$ (相差一个高级无穷小), 在此段时间内死亡的概率为 $q\mathrm{d}t$. 并设不同细菌的死亡、分裂为相互独立. 今设在时刻 t 有 n 个, 问在 $t+\mathrm{d}t$ 时将有多少? l 个分裂与 m 个死亡的概率为

$$\frac{n!}{l!m!(n-l-m)!}(p\mathrm{d}t)^l(q\mathrm{d}t)^m(1-(p+q)\mathrm{d}t)^{n-l-m},$$

除 (A) $l=0,m=0$, (B) $l=1,m=0$, (C) $l=0,m=1$ 三个情形外, 此概率都是 $\mathrm{d}t$ 的二阶以上的无穷小. $t+\mathrm{d}t$ 时的细菌数在 (A) 为 n, 在 (B) 为 $n+1$, 在 (C) 为 $n-1$. 故经 $\mathrm{d}t$ 时后由 n 转入 k 的概率 $P(\mathrm{d}t,n,k)$ 为 (相差一个高级无穷小)

$$\left.\begin{aligned}P(\mathrm{d}t,n,n)&=1-n(p+q)\mathrm{d}t,\\P(\mathrm{d}t,n,n+1)&=np\mathrm{d}t,\\P(\mathrm{d}t,n,n-1)&=nq\mathrm{d}t.\end{aligned}\right\}\tag{48.1}$$

若一度为 0, 则永久为 0, 故有

$$P(\mathrm{d}t,0,0)=1.\tag{48.2}$$

这样的随机过程称为生灭过程 (birth and death process). 若稍推广 (48.1) 而令

$$\left.\begin{aligned}P(\mathrm{d}t,n,n)&=1-(p_n+q_n)\mathrm{d}t,\\P(\mathrm{d}t,n,n+1)&=p_n\mathrm{d}t,\\P(\mathrm{d}t,n,n-1)&=q_n\mathrm{d}t,\end{aligned}\right\}\tag{48.1$'$}$$

也采用同样的名字. 例如, 随着 n 的增大, 在食粮不易找到, 分裂率降低, 死亡率提高的情形下, 自然地可设

$$p_1>\frac{p_2}{2}>\frac{p_3}{3}>\cdots,\quad q_1<\frac{q_2}{2}<\frac{q_3}{3}<\cdots.\tag{48.2$'$}$$

特别当 $q_n\equiv 0$ 时, 叫做**纯生过程** (pure birth process), 当 $p_n\equiv 0$ 时, 叫做**纯灭过程** (pure death process).

若用所述的 Markoff 过程理论来处理这一过程, 则存在着困难. 因为现在状态空间是 $\{0,1,2,3,\cdots\}$, 它不是紧致的. 简单的想法是, 若添上 ∞ 而再规定

$$P(\mathrm{d}t,\infty,\infty)=1$$

是否可以? 这样做未必可行, 如何克服此困难, 将在以后叙述.

暂不顾这困难, 与有限个状态时一样, 可如下考虑这一过程: 若自 n 出发, 经指数型逗留时间 τ(其平均值为 $(p_n+q_n)^{-1}$) 后, 转入 $n+1$ 或 $n-1$, 其概率分别为 $p_n/(p_n+q_n), q_n/(p_n+q_n)$. 然后从新的状态 $n-1$ 或 $n+1$ 同样地继续运动.

首先考虑纯灭过程. 以后 q_n 都设为正, 此时, 细菌数逐渐减少, 其个数最后为 0 而灭绝. 今试求**灭绝时间** (extinction time) 的分布, 这是 Markoff 时间, 但不深究. 自 n 出发, 按 $n \to n-1 \to n-2 \to \cdots \to 1 \to 0$ 进行, 若设停留于 k 的时间为 τ_k, 则灭绝时间 ε_n 为

$$\varepsilon_n = \tau_n + \tau_{n-1} + \cdots + \tau_1,$$

其中 $\tau_n, \tau_{n-1}, \cdots, \tau_1$ 都是独立的, 并且各服从于平均值为 $q_k^{-1}(k = n, n-1, \cdots, 1)$ 的指数分布 $F_k(t) = 1 - \mathrm{e}^{-q_k t}$, 故 ε_n 的分布就是其卷积. 若求 ε_n 的平均值, 则得

$$E(\varepsilon_n) = \sum_{k=1}^{n} E(\tau_k) = \sum_{k=1}^{n} q_k^{-1} < \infty,$$

因此 $P(\varepsilon_n < \infty) = 1$, 即不论从多么大的个数出发, 迟早总会灭绝.

为了决定 $P(t, n, k)$, 记分布 $F_k, \cdots, F_n$ 的卷积为 $F_{k,n}$. 这显然是 $\tau_n + \tau_{n-1} + \cdots + \tau_k$ 的分布, $F_{k,n}(t)$ 就是在时间 t 以前个数减到不大于 $k-1$ 的概率. 因此, 经过时间 t 后自 n 转移到 k 的概率为

$$P(t, n, k) = F_{k+1,n}(t) - F_{k,n}(t), \quad k = 0, 1, 2, \cdots, n-1, n,$$

这里

$$F_{0,n}(t) = 0, \quad F_{n+1,n}(t) = 1.$$

为了要使状态空间添上 ∞ 后紧化为 R, 除上述诸转移概率外, 还必须合理地定义 $P(t, \infty, k), k = 0, 1, 2, \cdots, \infty$. 为此, 对 R 上的连续函数 $f(n)$, 必须使

$$T_t f(n) = \sum p(t, n, k) f(k)$$

连续. 因为除 ∞ 外, R 中其他的点都是孤立点, 所谓连续就是要求在 ∞ 的极限值与实际值相等. 今特别取 $f(m) = \delta_{mk}$, 这函数对 $k \neq \infty$ 是连续的, 此时因 $T_t f(n) = P(t, n, k)$, 故必须要求

$$\lim_{n\to\infty} P(t, n, k) = P(t, \infty, k) \quad k \neq \infty.$$

为求此极限试先求 $\lim_{n\to\infty} F_{k,n}(t)$. $F_{k,n}(t)$ 是 $\tau_n^n + \tau_{n-1}^n + \cdots + \tau_k^n$ 的分布函数 (上标 n 表示自状态 n 出发), 而 $\tau_n^{n+1}, \tau_{n-1}^{n+1}, \cdots, \tau_k^{n+1}$ 分别与 $\tau_n^n, \tau_{n-1}^n, \cdots, \tau_k^n$ 有相同的分布, 故 $\tau_n^{n+1} + \tau_{n-1}^{n+1} + \cdots + \tau_k^{n+1}$ 也与 $\tau_n^n +$

$\tau_{n-1}^n + \cdots + \tau_k^n$ 同分布, 故得

$$\begin{aligned} F_{k,n+1}(t) &= P(\tau_{n+1}^{n+1} + \tau_n^{n+1} + \cdots + \tau_k^{n+1} \leqslant t) \\ &\leqslant P(\tau_n^{n+1} + \cdots + \tau_k^{n+1} \leqslant t) \\ &= P(\tau_n^n + \cdots + \tau_k^n \leqslant t) = F_{k,n}(t). \end{aligned}$$

故 $F_{k,n}(t)$ 是 n 的单调不增函数, 记其极限为 $G_k(t)(\geqslant 0)$, 得

$$P(t, \infty, k) = G_{k+1}(t) - G_k(t).$$

因此

$$\begin{aligned} P(t, \infty, \infty) &= 1 - \sum_{k \neq \infty} P(t, \infty, k) \\ &= 1 - \lim_{m \to \infty} \sum_{k=0}^{m} (G_{k+1}(t) - G_k(t)) = 1 - \lim_{m \to \infty} G_{m+1}(t). \end{aligned}$$

应区别两种情况:

(i) $\sum\limits_{k=1}^{\infty} q_k^{-1} = \infty$. 因 $F_{k,n}(t) \downarrow G_k(t)(\geqslant 0)$, 故对 $\lambda > 0$ 有

$$\begin{aligned} \int_0^\infty \mathrm{e}^{-\lambda t} G_k(t) \mathrm{d}t &= \lim_{n \to \infty} \int_0^\infty \mathrm{e}^{-\lambda t} F_{k,n}(t) \mathrm{d}t \\ &= \lim_{n \to \infty} \frac{1}{\lambda} \int_0^\infty \mathrm{e}^{-\lambda t} \mathrm{d}F_{k,n}(t). \end{aligned}$$

因为分布 $F_{k,n}(t)$ 是 $F_k, F_{k+1}, \cdots, F_n$ 的卷积, 故

$$\begin{aligned} &= \lim_{n \to \infty} \frac{1}{\lambda} \prod_{v=k}^{n} \int_0^\infty \mathrm{e}^{-\lambda t} \mathrm{d}F_v(t) \\ &= \lim_{n \to \infty} \frac{1}{\lambda} \prod_{v=k}^{n} \left(1 + \frac{\lambda}{q_k}\right)^{-1} = \frac{1}{\lambda} \prod_{v=k}^{\infty} \left(1 + \frac{\lambda}{q_k}\right)^{-1}. \end{aligned}$$

由于 $\sum q_k^{-1} = \infty$, 此无穷乘积等于 0. 又因 $G_k(t)$ 和 $F_{k,n}(t)$ 对 t 都是单调不减的, 故必有 $G_k(t) \equiv 0$. 于是

$$P(t, \infty, \infty) = 1, \quad 因此\ P(t, \infty, k) = 0, \quad k \neq \infty.$$

此时 ∞ 为套点, 而上述的简单想法可行.

(ii) $\sum\limits_1^{\infty} q_k^{-1} < \infty$. 此时上面的无穷乘积恒为正, 当 $k \uparrow \infty$ 时趋于 1, 故得

$$\lim_{k\to\infty}\int_0^{\infty} \mathrm{e}^{-\lambda t} G_k(t)\mathrm{d}t = \frac{1}{\lambda}.$$

由于 $G_k(t)$ 与 $F_{k,n}(t)$ 都对 t 单调不减, 故得

$$\int_0^{\infty} \mathrm{e}^{-\lambda t} \lim_{k\to\infty} G_k(t)\mathrm{d}t = \frac{1}{\lambda},$$

由此得

$$\lim_{k\to\infty} G_k(t) \equiv 1.$$

因此

$$P(t,\infty,\infty) = 0, \quad \sum_{k\neq\infty} P(t,\infty,k) = 1.$$

故 ∞ 是瞬时逗留点, 由 ∞ 到有限点的转移是重要的, 此时前述的简单想法不可行.

上述两种情况都不能从有限点到达无限大, 但在 (i) 中, 在 ∞ 附近的减少程度逐渐变弱, 故自然伸延后, ∞ 不能不成为套点. 反之, 在 (ii) 中, 则越接近 ∞, 减少程度越强烈, 故即使在 ∞ 也不能不回到有限点. 因为有限点为指数型逗留点, 故若以 R 的点为横轴, 时间为纵轴, 绘出自 ∞ 出发的样本过程则得图 48-1.

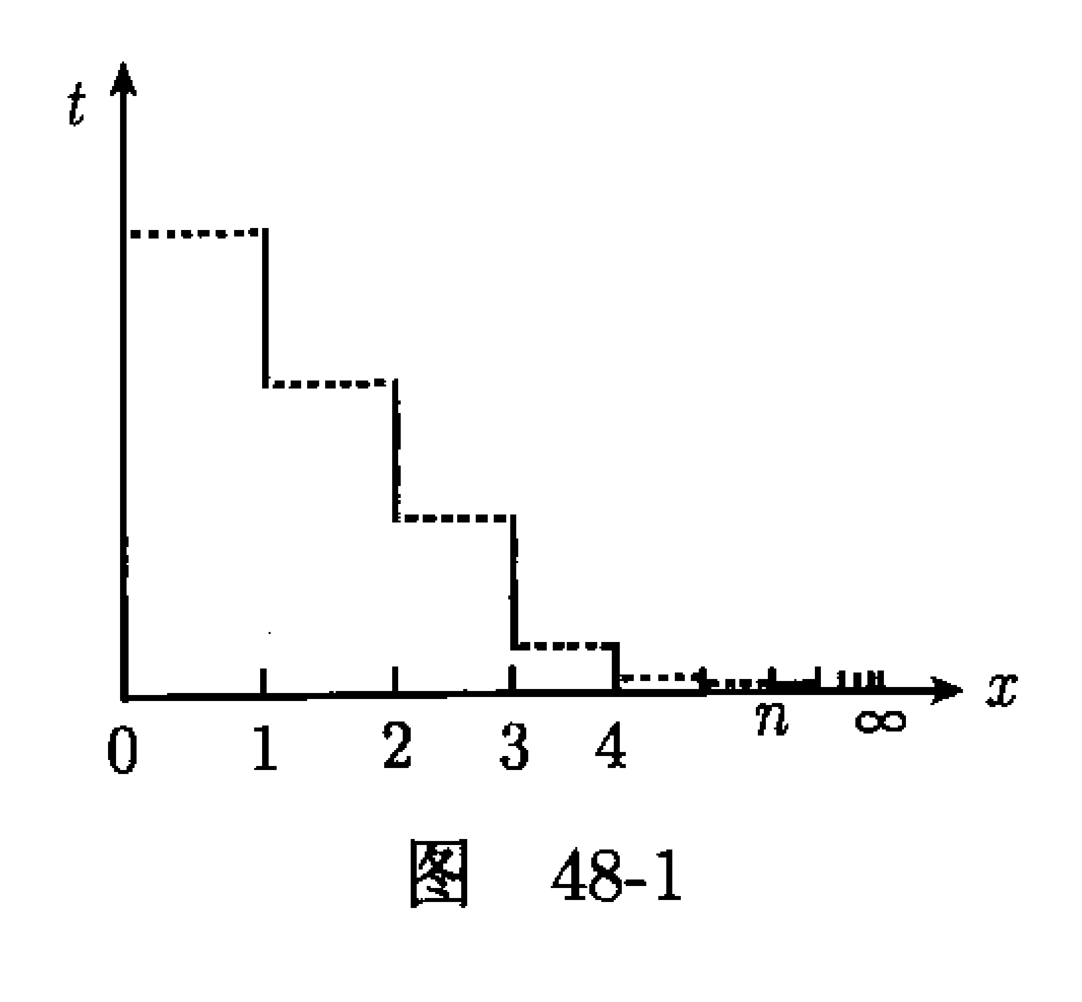


图　48-1

下面考虑纯生过程. 设 $p_n > 0 (n \geqslant 1)$. 0 当然是套点. 自 $n(\geqslant 1)$ 出发, 个数逐渐增加, 若以 τ_{nm} 表示从 n 出发增至 $m(> n)$ 所用的时间, 则与上同样

$$E(\tau_{nm}) = \sum_{n}^{m-1} p_{\nu}^{-1} < \infty.$$

故从 n 可以实际到达 m. 若令

$$\sum_1^{\infty} p_n^{-1} < \infty.$$

则自 1 出发 (因之也可自 n 出发) 在有限时间 τ 内, 可超过任何大的数. 因此对适当大的 t,

$$\sum_{m\neq\infty} P(t,n,m) < 1.$$

故若想在 $R=\{1,2,\cdots,\infty\}$ 上定义 Markoff 过程, 则应设

$$P(t,n,\infty) = 1 - \sum_{m\neq\infty} P(t,n,m).$$

因 $P(t,n,m)\equiv 0, m<n$, 故必须

$$P(t,\infty,m) = \lim_{n\to\infty} P(t,n,m) = 0.$$

亦即 ∞ 是套点. 以 τ_n 表示自有限点 n 到达 ∞ 所用的有限时间, 则

$$E(\tau_n) = \sum_{n}^{\infty} p_\nu^{-1} < \infty.$$

这 τ_n 是 Markoff 时间, 称为**爆发时间** (explosion time).

若令

$$\sum_{1}^{\infty} p_n^{-1} = \infty,$$

则

$$\begin{aligned} E\{\mathrm{e}^{-\lambda\tau_{nm}}\} &= \prod_{\nu=n}^{m-1}\int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{e}^{-tp_\nu}p_\nu \mathrm{d}t \\ &= \prod_{\nu=n}^{m-1}\left(1+\frac{\lambda}{p_\nu}\right)^{-1} \to 0 \quad (m\to\infty). \end{aligned}$$

令 $\tau_n = \lim_m \tau_{nm}$, 则 $E=(\mathrm{e}^{-\lambda\tau_n})=0, P(\tau_n=\infty)=1$, 即永远不会爆发, 而

$$\sum_{m\neq\infty} P(t,n,m) = 1.$$

与前同理可知 ∞ 是套点.

对一般的生灭过程, 由 p_n, q_n 的大小关系可产生各种有趣现象, 但从略.

第5章　扩　散

§49　扩 散 点

设 $x^{(a)}(t,\omega)$ 为在 R 内变动的 Markoff 过程, 记号及条件都同第 4 章. $x^{(a)}(t,\omega)$ 的样本过程只有第一类不连续点. 令 U 为 R 的开子集. 从 U 内的任意点 a 出发的 $x^{(a)}(t,\omega)$, 在 U 的最初通过时间 $\tau=\tau_U^{(a)}$ 以前 (包含 τ 在内), 如果以概率 1 为连续的, 则此 Markoff 过程称为在 U 内是**扩散的**. 由连续性假定, $x^{(a)}(\tau_U^{(a)})$ 位于 U 的边界 ∂U 上.

如果在 b 的某一邻域内 $x^{(a)}(t)$ 是扩散的, 则称 $b\in R$ 为扩散点. 扩散点的集合显然是开集.

一般地, 算子 S 称为在 x 是**局部的**(local), 就是指在 x 的邻域上, 对 $f=g$ 的 f, $g\in\mathfrak{D}(S)$, $Sf(x)=Sg(x)$ 成立.

定理 1　$x^{(a)}(t,\omega)$ 的生成算子 A 在此 Markoff 过程的扩散点是局部的.

证明　令 a 为扩散点. 若令 a 的扩散邻域为 U, 则包含于 U 内的 a 的邻域都是扩散的. 今设 f_1, $f_2\in\mathfrak{D}(A)$; 且在 a 的邻域 V 上 $f_1=f_2$. 由 Dynkin 定理可得

$$Af_i(a)=\lim_{W\downarrow a}\frac{Ef_i(x^{(a)}(\tau_W))-f_i(a)}{E(\tau_W)}$$

若取 W 使 $\overline{W}\subset U\cap V$, 则 $x^{(a)}(\tau_W)\in U\cap V$, 而且 $Af_1(a)=Af_2(a)$.

§50　Ray 定 理

考虑 R 上的 Markoff 过程 $x^{(a)}(t)$. R 的点 b 称为**一维点**, 就是指存在 b 的某一邻域与线段同胚. 这可以看成是只与 R 有关的性质, 而与 $x^{(a)}(t)$ 无关. 一维点全体显然是开集合. 若 b 是 R 的一维点, 同时又是 $x^{(a)}(t)$ 的扩散点, 则称 b 为 $x^{(a)}(t)$ 的**一维扩散点**. 这种点的全体也是开集合.

令 b 为扩散点, U 为 b 的邻域. 由于 b 是扩散点, 所以不难想象, 从 b 出发的 Markoff 过程 $x^{(b)}(t)$ 在充分短的时间内跑出 U 外的概率 $P(t,b,U^c)$ 非常小. 断定此概率为 $o(t)$ 就是本节所说明的 Ray 定理.

定理 1 若 b 是 $x^{(a)}(t)$ 的一维扩散点, 则

$$P(t,b,U^c)=o(t),\quad U\text{是 } b \text{ 的邻域}. \tag{50.1}$$

证明 由假定, b 附近的点都是一维扩散点. 因为 U 越小则 $P(t,b,U^c)$ 越大, 所以只对 U 的点都是一维扩散点的情形予以证明. 因为 b 有与线段同胚的邻域, 故 b 的某邻域内的点可以由对应的线段上的点即实数来表示. U 也可看成区间 (u_1,u_2).

若 (50.1) 不成立, 则存在 $t_n\downarrow 0$, $c>0$ 使

$$p(t_n,b,U^c)>ct_n. \tag{50.2}$$

现在令 $\tau_1^{(b)}$ 为 $x^{(a)}(t,\omega)$ 初次从 u_1 跑出 U 的时间. 若 $x^{(a)}(t,\omega)$ 永远停留于 U 内, 或从 u_2 跑出 U, 则定义 $\tau_1^{(b)}=\infty$. 精确些说就是: 令 U 的最初通过时间为 $\tau_U^{(b)}$, 若 $\tau_U^{(b)}<\infty$, 则因 U 只由扩散点所构成, 故 $x^{(b)}(\tau_U^{(b)})$ 等于 U 的边界点 u_1 或 u_2. 定义

$$\tau_1^{(b)}=\begin{cases}\tau_U^{(b)}, & \text{当}\quad x^{(b)}(\tau_U^{(b)})=u_1\text{时};\\ \infty, & \text{其他}.\end{cases}$$

以 u_2 代替 u_1, 则可同样定义 $\tau_2^{(b)}$. 由 (50.2) 有

$$P(\tau_1^{(b)}\leqslant t_n)+P(\tau_2^{(b)}\leqslant t_n)\geqslant P(x^{(b)}(t_n)\in U^c)>ct_n.$$

所以或者

$$P(\tau_1^{(b)}\leqslant t_n)>\frac{c}{2}t_n$$

对无穷多个 n 成立; 或者

$$P(\tau_2^{(b)}\leqslant t_n)>\frac{c}{2}t_n$$

对无穷多个 n 成立. 试证两者之一成立而由它导出矛盾. 因为证法相同, 所以只证后者. 改用记号而将自下式导出矛盾:

$$P(\tau_2^{(b)}\leqslant t_n)>ct_n,\quad c>0,\quad t_n\downarrow 0. \tag{50.3}$$

若在 b 与 u_2 之间取 y, 则

$$P(\tau_2^{(b)} \leqslant t_n) \leqslant P(\tau_2^{(b)} - \tau_2^{(b)}(y) \leqslant t_n, \tau_2^{(b)}(y) \leqslant t_n).$$

这里 $\tau_2^{(b)}(y)$ 是以 (u_1, y) 代上面的 U 后定义的 $\tau_2^{(b)}$.

由强 Markoff 性

$$\begin{aligned}&=E\{P(\tau_2^{(b)} - \tau_2^{(b)}(y) \leqslant t_n/B^{(b)}_{\tau_2}(y)); \tau_2^{(b)}(y) \leqslant t_n\}\\&=E\{P(\tau_2^{(y)} \leqslant t_n); \tau_2^{(b)}(y) \leqslant t_n\} \leqslant P(\tau_2^{(y)} \leqslant t_n).\end{aligned}$$

($\tau_2^{(y)}$ 与 $\tau_2^{(b)}$ 类似, 同样地对 $U=(u_1,u_2)$ 定义, 只是以 y 代替 b.) 因而由 (50.3), 对 $b \leqslant y < u_2$ 得

$$P(\tau_2^{(y)} \leqslant t_n) > ct_n, \quad c>0, \quad t_n \downarrow 0. \tag{50.4}$$

若令 $\varepsilon = (u_2 - b)/4 (>0), a = b + 2\varepsilon$ 及

$$y(t,\omega) = \begin{cases} x^{(a)}(t,\omega), & t < \tau_U^{(a)}, \\ x^{(a)}(\tau_U^{(a)}, \omega), & t \geqslant \tau_U^{(a)}, \end{cases}$$

则 $y(t)$ 的样本过程是连续的, 且恒处于 $[u_1, u_2]$ 内. 故若取 s 充分小, 则

$$P(a-\varepsilon < y(t) < a+\varepsilon, 0 \leqslant t \leqslant s) > \frac{1}{2}. \tag{50.5}$$

因为 $y(t)$ 的样本过程在 $0 \leqslant t \leqslant s$ 上一致连续, 故

$$\alpha_n = P(\text{对某一 } k(kt_n \leqslant s), y((k-1)t_n) < a+\varepsilon, y(kt_n) = u_2) \to 0 \quad (n \to \infty). \tag{50.6}$$

然而

$$\alpha_n \geqslant \sum_{k=1}^{[s/t_n]} P(a-\varepsilon < y(t) < a+\varepsilon, 0 \leqslant t \leqslant (k-1)t_n, \text{而且 } y(kt_n) = u_2).$$

因 $y(t)$ 关于 $\boldsymbol{B}_t$ 可测, 故

$$\begin{aligned}&= \sum_{k=1}^{[s/t_n]} E\{P(y(t_n) = u_2/\boldsymbol{B}_{(k-1)t_n}); \quad a-\varepsilon < y(t) < a+\varepsilon, 0 \leqslant t \leqslant (k-1)t_n\}\\&= \sum E\{P(\tau_2^{(b)} \leqslant t_n)_{b=x^{(a)}((k-1)t_n)}; \quad a-\varepsilon < y(t) < a+\varepsilon, 0 \leqslant t \leqslant (k-1)t_n\}.\end{aligned}$$

$x^{(a)}((k-1)t_n)$ 在 $\{a-\varepsilon<y(t)<a+\varepsilon, 0\leqslant t\leqslant(k-1)t_n\}$ 上等于 $y((k-1)t_n)$, 因而处于 $(a-\varepsilon,a+\varepsilon)$ 内. 因 $(a-\varepsilon,a+\varepsilon)\subset[b,u_2)$, 故由 (50.4) 知 $P(\tau_2^{(b)}\leqslant t_n)_{b=x^{(a)}((k-1)t_n)}$ 大于 ct_n.

$$\begin{aligned}\alpha_n&\geqslant\sum_k ct_nP(a-\varepsilon<y(t)<a+\varepsilon,0\leqslant t\leqslant(k-1)t_n)\\&\geqslant\frac{1}{2}ct_n[s/t_n]\geqslant\frac{c}{4}s\quad(n\geqslant 2).\end{aligned}$$

此与 (50.6) 矛盾.

定理 2 设 R 的开集 U 与线段同胚, 并设 U 所有的点都是扩散点. 若令 F 为 U 的闭子集, 则对 $b\in F$ 一致地有

$$P(t,b,U^c)=o(t). \tag{50.7}$$

证明 如果扩大定理中 F 而缩小 U 时, 定理能被证明, 则此定理自然就成立了. 故不妨设 $\bar{U}$ 的各点也都是一维扩散点, 而且 F 与闭线 J 同胚, 令 $J=[b_1,b_2]$. 由定理 1

$$P(t,b_i,U^c)=o(t),\quad i=1,2. \tag{50.8}$$

此处 $o(t)$ 与 i 无关. 利用强 Markoff 性, 对 $b_1\leqslant b\leqslant b_2$,

$$P(t,b,U^c)=\int_0^t\varphi_1(\mathrm{d}s)P(t-s,b_1,U^c)+\int_0^t\varphi_2(\mathrm{d}s)P(t-s,b_2,U^c)=o(t).$$

这里 φ_i 是 $x^{(b)}(t)$ 从 b_i 跑出 (b_1,b_2) 的时间的分布.

回顾定理 1 的证明, 可见

$$P(\tau_1^{(b)}<t)+P(\tau_2^{(b)}<t)=o(t),$$

故如以 $Q(t,b,U^c)$ 表 $x^{(b)}(s)$ 在时间区间 $0\leqslant s\leqslant t$ 内到过 U^c 的概率, 则上式左边为 $Q(t,b,U^c)$, 故得下面的定理.

定理 3 定理 1 和定理 2 对 $Q(t,b,U^c)$ 也成立.

§51 局部生成算子

称 Markoff 过程 $x^{(a)}(t)$ 在点 b 上有局部性, 如对 b 的任意邻域 U, 有

$$P\{x^{(b)}(t)\in U^c\}=P(t,b,U^c)=o(t). \tag{51.1}$$

根据上节的 Ray 定理, $x^{(a)}(t)$ 在其一维扩散点上有局部性.

以下设在 b 有局部性. $f(x)$ 在 $x=b$ 的充分小的邻域 V 上连续而且有界时, 若

$$\lim_{t\downarrow 0}\frac{1}{t}\left\{\int_V P(t,b,\mathrm{d}y)f(y)-f(b)\right\} \tag{51.2}$$

存在, 则说 $f\in\mathfrak{D}(A_b)$, 并以 A_bf 表示此极限. 因为在 b 上有局部性, 故此定义与 V 的选择无关. A_b 有下列性质.

(A$_\mathrm{b}$.1) (局部性) 若 $f\in\mathfrak{D}(A_b)$ 且在 b 的附近 (即某邻域上)$f=g$, 则 $g\in\mathfrak{D}(A_b)$ 且

$$A_bf=A_bg.$$

(A$_\mathrm{b}$.2) (线性) 若 $f,g\in\mathfrak{D}(A_b)$, 则 $\alpha f+\beta g\in\mathfrak{D}(A_b)$, 且

$$A_b(\alpha f+\beta g)=\alpha A_bf+\beta A_bg.$$

(A$_\mathrm{b}$.3) (正定性) 若在 b 的邻域上 $f\geqslant f(b), f\in\mathfrak{D}(A_b)$, 则 $A_bf\geqslant 0$.

(A$_\mathrm{b}$.4) 若 $f\in\mathfrak{D}(A)$, 则 $f\in\mathfrak{D}(A_b)$, 且 $A_bf=Af(b)$.

有了这些准备知识, 就可以定义局部生成算子 A_u 了.

定义 1 设开集 U 中各点均有局部性, 由

$$\mathfrak{D}(A_U)=\left\{f\middle/\begin{array}{l}\text{对所有的}b\in U, f\in\mathfrak{D}(A_b),\\ A_bf\text{对}b\in U\text{连续}\end{array}\right\}$$
$$A_Uf(b)=A_bf, b\in U$$

定义的算子 A_U 称为 $x^{(a)}(t)$ 在 U 上的**局部生成算子**(local generator).

前面已经看到, 一维扩散点上有局部性, 故在其上可以考虑 A_b, 且这时可把 A_b 如下定义.

定义 2 $P_U(t,a,E)=P\{\omega/x^{(a)}(t)\in E$, 而且 $x^{(a)}(s)\in U, 0\leqslant s\leqslant t\}$

$$A_bf=\lim_{t\downarrow 0}\frac{1}{t}\left\{\int_U P_U(t,b,\mathrm{d}y)f(y)-f(b)\right\}.$$

这个定义比较有用.

两个定义的一致是由于

$$\left|\int_U P(t,b,\mathrm{d}y)f(y)-\int_U P_U(t,b,\mathrm{d}y)f(y)\right|$$

$$\leqslant \int_U (P(t,b,\mathrm{d}y) - P_U(t,b,\mathrm{d}y))|f(y)|$$
$$\leqslant \sup_U |f| \cdot P\{对于0 \leqslant s \leqslant t, x^{(a)}(s) \notin U\}.$$

由 §50 的定理 3, 后一概率是 $o(t)$.

§52 一维扩散点的分类

一维扩散点的全体是开集合, 而且可以表为至多可数多个连续分量的直和, 并且各分量与实数开区间同胚. 今取其中一个 I, 若 I 与 (r_1,r_2) 同胚, 则 I 的点可由 (r_1,r_2) 中的实数表示.

令 $b\in I$, 则 $x^{(b)}(t)$ 在 $0\leqslant t<\tau_I^{(b)}$ 中连续的概率为 1. 若

$$P(对0\leqslant t<\tau_I^{(b)}中的t, x^{(b)}(t)\geqslant b)=1, \tag{52.1}$$

则称 b 为**右通过点**(right translation point).

令 U 为含于 I 内的 b 的邻域. 因恒有 $\tau_U^{(b)}\leqslant\tau_I^{(b)}$, 故若 b 是右通过点, 则

$$P(对0\leqslant t<\tau_U^{(b)}中的t, x^{(b)}(t)\geqslant b)=1. \tag{52.2}$$

反之, 由此条件可知 b 是右通过点. 为了证明这一点, 令 τ 为 $x^{(b)}(t)$ 自 b 跑出 $[b,r_2]$ 的最初时间. 如此情况不发生, 则令 $\tau=\infty$, τ 是 Markoff 时间. 令 $\tau_n=\min(\tau,n)$, 则 τ_n 是有限 Markoff 时间. 需要证明的是 $P(\tau<\infty)=0$. 因为

$$P(\tau<\infty)=\lim_{n\to\infty}P(\tau<n)=\lim_{n\to\infty}P(\tau_n<n),$$

故只要证明 $P(\tau_n<n)=0$ 即可. 由 τ 的定义,

$$\begin{aligned}
0 =&P(\tau<n, 且对充分小的t, x^{(b)}(\tau_n+t)\geqslant b)\\
=&P(\tau_n<n, 且对充分小的t, x^{(b)}(\tau_n+t)\geqslant b)\\
=&E\{P\{对充分小的t, x(\tau_n+t)\geqslant b/\boldsymbol{B}_{\tau_n}\};\tau_n<n\}\\
=&E\{P\{对充分小的t, x^{(b)}(t)\geqslant b\};\tau_n<n\}\\
\geqslant&E\{P\{对0\leqslant t<\tau_U^{(b)}, x^{(b)}(t)\geqslant b\};\tau_n<n\}\\
=&P\{\tau_n<n\}\quad (由假定(52.2)).
\end{aligned}$$

在 (52.1) 中, 以 $x^{(b)}(t) \leqslant b$ 代替 $x^{(b)}(t) \geqslant b$, 即得**左通过点**(left translation point) 的定义. 与 (52.2) 同样, 在此定义中 I 也可由 b 的任意邻域 $U(\subset I)$ 来代替.

既是左通过点又是右通过点的点, 实际上就是套点. 是右通过点但不是套点的点, 称为**纯右通过点**. 同样定义**纯左通过点**. I 中既非左通过点又非右通过点的点叫做**正则点**(regular point).

左通过点、右通过点、纯左通过点、纯右通过点、正则点与套点各自的全体分别记为 $\Lambda_l, \Lambda_r, \Lambda_{pl}, \Lambda_{pr}$, Λ_2 与 Λ_t.

虽然 I 与实数开区间同胚, 但 I 与此区间的同胚对应却有好几种, 今按方向把它分为两种. 取 I 内任两点, 按其所对应的实数大小关系而分为两种. 这样的分类与两点的选择无关. 关于同向的两个表示, 左通过点和右通过点的定义是一致的, 但关于逆向的表示, 则需要左右互换. 套点和正则点的定义则对任何表示都一样.

(i) 若令 b 为右通过点, 则 $x^{(a)}(t)$ 从右向左切 b 的概率为 0.

证明　若令 τ 为从右向左最初切 b 的时间, 则 τ 是 Markoff 时间 (若不发生这种情况, 则令 $\tau = \infty$). 用自 (52.2) 导出 (52.1) 同样的方法, 可证明 $P(\tau < \infty) = 0$.

(ii) 对 $b < a < r_2$ 内所有的 a, 若

$$P\{\text{对}0 \leqslant t < \tau_I^{(a)}\text{中的}t, x^{(a)}(t) \geqslant b\} = 1, \tag{52.3}$$

则 b 是右通过点.

由假定, 如果已给 t, 则对 $b < a < r_2$,

$$P\{\tau_I^{(a)} \leqslant t, \text{或}b \leqslant x^{(a)}(t) < r_2\} = 1. \tag{52.4}$$

在 b 与 r_2 之间取 u_2, 在 r_1 与 b 之间取 v_1, 在 b 与 u_2 之间取 v_2, 由 Ray 定理, 存在随 t 而收敛于 0 的 $\delta(t)$ 使得

$$P\{\text{对}0 \leqslant s < t, x^{(a)}(s) \in (r_1, u_2)\} > 1 - \delta(t) \cdot t.$$

这里可对 $v_1 \leqslant a \leqslant v_2$ 中的 a 共同取 $\delta(t)$. 特别对 $b < a \leqslant v_2$ 中的 a, 由假定 (52.3) 得

$$P\{\text{对}0 \leqslant s < t, x^{(a)}(s) \in [b, u_2)\} > 1 - \delta(t) \cdot t.$$

因此

$$P\{x^{(a)}(t) \in [b, u_2)\} > 1 - \delta(t) \cdot t,$$

即 $$P(t, a, [b, u_2)) > 1 - \delta(t) \cdot t.$$

当 $a \downarrow b$ 时, 因 $P(t, a, \cdot) \to P(t, b, \cdot)$(弱收敛), 故

$$P(t, b, [b, r_2)) \geqslant 1 - \delta(t) \cdot t,$$

即 $$P\{x^{(b)}(t) \in [b, r_2)\} \geqslant 1 - \delta(t) \cdot t.$$

故

$$P\{\tau_I^{(b)} \leqslant t, \text{或} b \leqslant x^{(b)}(t) < r_2\} \geqslant 1 - \delta(t) \cdot t. \tag{52.5}$$

又由 (52.4), 对 $b < a < r_2$ 中的 a 有

$$P\{\tau_I^{(a)} \leqslant t, \text{或} b \leqslant x^{(a)}(t) < r_2\} \geqslant 1 - \delta(t) \cdot t. \tag{52.6}$$

若对任意给定的 t, 能证明

$$P\{\text{对} 0 \leqslant s < \min(\tau_I^{(b)}, t) \text{中的} s, x^{(b)}(s) \in [b, r_2)\} = 1, \tag{52.7}$$

则令 $t \uparrow \infty$ 即知 b 是右通过点. 为此, 若考虑到 $x^{(b)}(s)$ 在 $0 \leqslant s < \tau_I^{(b)}$ 中连续, 则只要证明

$$P\left\{\text{对} \frac{k}{n}t < \tau_I^{(b)} \text{中的} k (0 \leqslant k \leqslant n), x^{(b)}\left(\frac{k}{n}t\right) \in [b, r_2)\right\} \to 1 \quad (n \to \infty)$$

即可. 利用 Markoff 性及 (52.5) 和 (52.6), 得

$$\text{上述概率} \geqslant \left(1 - \delta\left(\frac{t}{n}\right) \cdot \frac{t}{n}\right)^n > 1 - \delta\left(\frac{t}{n}\right) \cdot t \to 1.$$

(iii) Λ_r 是闭集合.

证明 设 $b_n \in \Lambda_r, b_n \to b$, 只需证 $b \in \Lambda_r$. 不失普遍性, 可设 $b_n \uparrow b$ 或 $b_n \downarrow b$.

当 $b_n \uparrow b$ 时, 由 $b_n \in \Lambda_r$ 及 (i) 得

$$P\{\text{对} 0 \leqslant t < \tau_I^{(b)} \text{中的} t, x^{(b)}(t) \in [b_n, r_2)\} = 1.$$

令 $n \to \infty$ 得

$$P\{对0 \leqslant t < \tau_I^{(b)}中的t, x^{(b)}(t) \in [b, r_2)\} = 1,$$

这表明

$$b \in \Lambda_r.$$

当 $b_n \downarrow b$ 时, 由 $b_n \in \Lambda_r$ 及 (i), 对满足 $a > b_n$ 的 a,

$$P\{对0 \leqslant t < \tau_I^{(a)}中的t, x^{(a)}(t) \in [b_n, r_2)\} = 1.$$

故

$$P\{对0 \leqslant t < \tau_I^{(a)}中的t, x^{(b)}(t) \in [b, r_2)\} = 1.$$

因为 $b_n \downarrow b$, 故上式对任意的 $a > b$ 成立. 由 (ii) 得 $b \in \Lambda_r$.

(iv) Λ_l 也是闭集, 因此 Λ_t 也是闭集, 但 Λ_2 是开集.

§53 Feller 典范尺度

由上节, 一维扩散点的集合是开集, 而且它的各个分量与实数区间同胚, 因此, 分量内的点可由区间 (r_1, r_2) 内的点来表示. 若只考虑其中正则点, 则又得 (r_1, r_2) 中的开子集, 因而也是至多可数多个区间的直和. 取其中一个区间或其中子区间, 表以 $I = (i_1, i_2)$. (i_1, i_2) 内的点都是正则点, 称它为**正则区间**(regular interval).

取 I 的子区间 $J = (j_1, j_2)$ 使 $\overline{J} \subset I$. 包括端点在内, J 仅含正则点. 取 $a \in \overline{J}, x^{(a)}(t)$ 自 a 出发, 在到达 j_1 之前先到 j_2 的概率记为 $s(a; j_2, j_1)$, 在到达 j_2 之前先到达 j_1 的概率记为 $s(a; j_1, j_2)$, 则

$$1 - s(a; j_2, j_1) - s(a; j_1, j_2)$$

是永停留于 J 内的概率. 下面的定理断定, 此概率为 0.

定理 1

(A) a 从 j_1 上升到 j_2 时, $s(a; j_2, j_1)$ 从 0 连续地上升 (狭义) 到 1.

(B) $s(a; j_2, j_1) + s(a; j_1, j_2) = 1$.

证明 分为五步.

(i) 当 $j_1 \leqslant a < b \leqslant j_2$ 时,

$$s(a; j_2, j_1) = s(a; b, j_1)s(b; j_2, j_1). \tag{53.1}$$

$s(a;j_2,j_1)$ 是从 a 出发, 在到达 j_1 前先到达 j_2 的概率, 则这种情况的发生应是: 从 a 出发, 在到达 j_1 前先到达 b, 其次从 b 出发, 在到达 j_1 前先到达 j_2. 考虑 $x^{(a)}(t)$ 在到达 j_1 前先到达 b 的时间 τ(若不发生这种情形, 则令 $\tau=\infty$), 因为这是 Markoff 时间, 故利用强 Markoff 性, 就得 (53.1).

(ii) 若 $j_1<a<j_2$, 则

$$0<s(a;j_2,j_1)<1. \tag{53.2}$$

为证此, 先设有满足 $s(a;j_2,j_1)=0$ $(j_1<a<j_2)$ 的 a 存在, 将由此导致矛盾. 令在 $a<b\leqslant j_2$ 中且使 $s(a;b,j_1)=0$ 的 b 的全体为 B. 显然 $j_2\in B$. 令 B 的下确界为 b_0, 取 $j_1<a'<b_0, b\in B$. 若 $a'\leqslant a$, 则由 (i),

$$s(a';b,j_1)=s(a';a,j_1)s(a;b,j_1)=0.$$

若 $a<a'<b_0$, 则

$$0=s(a;b,j_1)=s(a;a',j_1)s(a';b,j_1).$$

但由 $a'<b_0$ 有 $s(a;a',j_1)>0$, 故 $s(a';b,j_1)=0$. 因此, 无论如何, 总有 $s(a';b,j_1)=0$. 这说明

$$P\{\text{对}0\leqslant t<\tau_J^{(a')}, x^{(a')}(t)\leqslant b\}=1.$$

在 B 中取 b 使之逼近 b_0 而得

$$P\{\text{对}0\leqslant t<\tau_J^{(a')}, x^{(a')}(t)\leqslant b_0\}=1 \quad (j_1<a'<b_0).$$

由上节 (ii), b_0 是左通过点 (上节中虽只就右通过点作了叙述, 但对左通过点也一样). 因为 J 的点都是正则点, 故产生矛盾. 因此 $s(a;j_2,j_1)>0$. 同理 $s(a;j_1,j_2)>0$. 故

$$s(a;j_2,j_1)\leqslant 1-s(a;j_1,j_2)<1.$$

(iii) 由 (i) 与 (ii) 得,

$$\text{当}j_1\leqslant a<b\leqslant j_2\text{时}, s(a;j_2,j_1)<s(b;j_2,j_1).$$

这样, 完成了 (A) 中除连续性外的证明.

(iv) 下面证 (B). $\alpha=1-s(a;j_2,j_1)-s(a;j_1,j_2)$ 是 $x^{(a)}(t)$ 永停留于 J 内的概率, 要证 $\alpha=0$. 令 U 为开集合, 若对 $a\in U, x^{(a)}(t)$ 永停留于 U

内的概率为 0, 则称 U 为**发散的**. 若 $U \supset V$ 而且 U 是发散的, 则 V 也是发散的. 因为 $a \in J$ 不是套点, 故由 §44 的(xi), 对 a 的充分小的邻域 U 有 $E(\tau_U^{(a)}) < \infty$, 故 $P(\tau_U^{(a)} < \infty) = 1$, 因此 U 是发散的. 再证: 如 I 的两个开区间 $U_1 = (u_1, v_1)$ 及 $U_2 = (u_2, v_2), (u_1 < u_2 < v_1 < v_2)$, 都是发散的, 则其和 $U = (u_1, v_2)$ 也是发散的. 若令 $a \in (u_1, v_1)$, 则 $x^{(a)}(t)$ 的轨迹中,

$$a \to u_1, \quad a \to v_1 \to v_2, \quad a \to v_1 \to u_2 \to u_1,$$
$$a \to v_1 \to u_2 \to v_1 \to v_2, \quad \cdots$$

是跑出 U 外的情形, 而永停留于 U 的情形只是

$$a \to v_1 \to u_2 \to v_1 \to u_2 \to v_1 \to u_2 \to \cdots$$

(因为 U_1, U_2 是发散的, 故无其他情形). 永停留于 U 内的概率为

$$s(a; v_1, u_1) \cdot s(v_1; u_2, v_2) \cdot s(u_2; v_1, u_1) \cdot s(v_1; u_2, v_2) \cdot s(u_2; v_1, u_1) \cdots,$$

因 $s(v_1; u_2, v_2) < 1$, 故此无限积为 0.

如 $a \in (v_1, v_2)$, 也同样得证.

若以有限个发散的区间覆盖 $\bar{J}$(Borel-Lebesgue 覆盖定理 1), 再利用上述结果, 则 $\bar{J}$ 被一个发散的开区间覆盖. 故 J 也是发散的. 这说明 $\alpha = 0$.

(v) 最后证明 (A) 中的连续性. 先证

$$\lim_{b \uparrow j_2} s(a; b, j_1) = s(a; j_2, j_1). \tag{53.3}$$

由 (iv) 所证, $\tau_J^{(a)}$ 有限的概率为 1. 若 $x(\tau_J^{(a)}) = j_1$, 则 $x^{(a)}(t)$ 在 $0 \leqslant t \leqslant \tau_J^{(a)}$ 中取最大值, 且此值小于 j_2. 故若取充分接近于 j_2 的 b, 则 $x^{(a)}(t)$ 在到达 b 之前先到达 j_1, 因此, 取 $b_n \uparrow j_2$, 如对任一个 $b_n, x^{(a)}(t)$ 在到达 j_1 前先到达 b_n, 则实际上是在到达 j_1 前先到达 j_2, 故

$$\lim_{n \to \infty} s(a; b_n, j_1) = s(a; j_2, j_1).$$

又因 $s(a; b, j_1)$ 随 b 增大而减少 (因 $b < b'$ 时, $s(a; b', j_1) = s(a; b, j_1) s(b; b', j_1)$), 故 (53.3) 得证.

由 (i) 及 (53.3) 得

$$s(b; j_2, j_1) = \frac{s(a; j_2, j_1)}{s(a; b, j_1)} \to 1 \quad (b \uparrow j_2).$$

故 $s(b;j_2,j_1)$ 在 $b=j_2$ 时连续. 同理 $s(b;j_1,j_2)$ 在 $b=j_1$ 时也连续, 因而 $s(b;j_2,j_1)$ 在 $b=j_1$ 时连续. 由此得

$s(a;j_2,j_1)=s(a;b,j_1)\cdot s(b;j_2,j_1)\to s(b;j_2,j_1), a\uparrow b$时.

$s(a;j_1,j_2)\to s(b;j_1,j_2), a\downarrow b$时.

故当 $a\downarrow b$ 时,

$$s(a;j_2,j_1)=1-s(a;j_1,j_2)\to 1-s(b;j_1,j_2)=s(b;j_2,j_1).$$

故 $s(a;j_1,j_2)$ 关于 a 连续.

定理 2 若 $(j_1,j_2)\subset(k_1,k_2)$, 则对 $a\in[j_1,j_2]$,

$$s(a;k_2,k_1)=s(a;j_1,j_2)\cdot s(j_1;k_2,k_1)+s(a;j_2,j_1)\cdot s(j_2;k_2,k_1).$$

证明 对 Markoff 时间 $\tau^{(a)}_{(j_1,j_2)}$ 应用强 Markoff 性即可.

在此定理中, 令 $s(a;j_1,j_2)=1-s(a;j_2,j_1)$, 得

$$s(a;k_2,k_1)=\alpha s(a;j_2,j_1)+\beta,$$

这里 α,β 是与区间 (j_1,j_2) 和 (k_1,k_2) 有关的常数. 利用此结果, 可得下述基本定理.

定理 3 在 $I=(i_1,i_2)$ 内, 除线性关系外, 存在唯一的连续狭义增函数 $s(a)$, 使对任意的 $J=(j_1,j_2)$, $\overline{J}\subset I$, 有

$$s(a;j_2,j_1)=\frac{s(a)-s(j_1)}{s(j_2)-s(j_1)}.$$

证明 先定义一个 $s(a)$, 固定 $J^0=(j_1^0,j_2^0),\overline{J^0}\subset I$, 取包含 a 的任意的 $J=(j_1,j_2)$, $I\supset\overline{J}\supset\overline{J^0}$, 并令

$$s(a)=\alpha s(a;j_2,j_1)+\beta,$$

其中 α,β 由

$$s(j_2^0)=1,\quad s(j_1^0)=0$$

决定. 今取更大的 $J'=(j_1',j_2')$ 代替 J. 同样定出 $s(a)$, 并记为 $s'(a)$. 由定理 2, 对 $b\in(j_1,j_2), s(b;j_2',j_1')$ 与 $s(b;j_2,j_1)$ 是线性关系, 故 $s'(b)$ 与 $s(b)$ 也是线性关系. 又因 $s'(j_2^0)=s(j_2^0)=1, s'(j_1^0)=s(j_1^0)=0$, 故 $s'(b)\equiv s(b)(b\in(j_1,j_2))$. 特别令 $b=a$ 而得 $s'(a)=s(a)$. 这说明 $s(a)$ 的

决定与 J 的选择无关. 又由 $s(a)$ 的定义方法看出, $s(a;j_2,j_1)$ 是 $s(a)$ 的线性式, 且由 $s(j_2;j_2,j_1)=1, s(j_1;j_2,j_1)=0$ 得证定理中所需的关系.

反之, 如满足定理要求的 $s(a)$ 有两个为 $s_1(a), s_2(a)$, 则在任意的 $J(\overline{J}\subset I)$ 内, 它们有线性关系. 因为线性关系的系数只由在两个 a 上的值所决定, 故此系数与 J 的选择无关.

此定理中的 $s(a)$ 称为 I 内的**典范尺度**(canonical scale). 这有 (内在的) 概率意义, 且与 I 的坐标选择 (与实数区间的拓扑对应的方法) 无关. 典范尺度与下节中叙述的典范测度都是 W. Feller 提出的.

例　令 $x^{(a)}(t)$ 是 $R=\mathbb{R}^1\cup\{\infty\}$ 上的 Wiener 过程. 若令 $I=\mathbb{R}^1$, 则 I 的点都是正则点. 试证 $s(x)\equiv\alpha x+\beta$. 由 Wiener 过程的左、右对称性得

$$s\Big(\frac{j_1+j_2}{2};j_2,j_1\Big)=\frac{1}{2}.$$

因此

$$s\Big(\frac{j_1+j_2}{2}\Big)=\frac{1}{2}(s(j_1)+s(j_2)).$$

因为 $s(x)$ 是连续的, 故可知 $s(x)\equiv\alpha x+\beta$.

§54　Feller 典范测度

先粗略说明定义典范测度的过程. 与上节一样, 令 I 为正则区间. 取区间 J 使 $\overline{J}\subset I$, 且令

$$p_J(a)=E(\tau_J^{(a)}),\quad q_J(a)=-p_J(a).$$

可证 $q_J(a)$ 关于 $s(a)$ 是凸的,

$$m_J(a)=\frac{\mathrm{d}q_J(a)}{\mathrm{d}s(a)}$$

是 a 的增函数. $m_J(a)$ 不一定连续. 若令 $J\subset J'$, 则对 $a\in J$, 可得

$$m_J(a)=m_{J'}(a)+\text{常数},$$

故若在 I 内任取一定点 a_0, 按 $m_J(a_0)=0$ 而规范化, 则得

$$m_J(a)=m_{J'}(a),$$

因此, 把它定义为 $m(a)$ 就得 I 上的增函数. 测度 $\mathrm{d}m$ 就是 Feller**典范测度**(canonical measre). s 一经决定, 则 m 除附加常数外可以确定, 若令 s 为 $s' = \alpha s + \beta$, 则 $m = \alpha^{-1} m + \beta'$. 需要证明的是下面的三个定理.

定理 1 $p_J(a) < \infty$.

证明 与上节证明 J 发散一样, 考虑 $u_1 < u_2 < v_1 < v_2$, 对 $J_1 = (u_1, v_1), J_2 = (u_2, v_2)$, 假定 $p_{J_1}(a), p_{J_2}(a) < \infty$, 然后对 $J = (u_1, v_2)$ 导出 $p_J(a) < \infty$ 即可. 因为已知 $x^{(a)}(t)$ 跑出 J 的概率为 1, 故可就轨迹计算 $p_J(a) = E(\tau_J^{(a)})$. 利用强 Markoff 性, 当 $a \in J_1$ 时,

$$\begin{aligned} p_J(a) =& p_{J_1}(a) + s(a; v_1, u_1) p_{J_2}(v_1) + s(a; v_1, u_1) s(v_1; u_2, v_2) p_{J_1}(u_2) \\ &+ s(a; v_1, u_1) s(v_1; u_2, v_2) s(u_2; v_1, u_1) p_{J_2}(v_1) + \cdots, \end{aligned}$$

由于 $s(v_1; u_2, v_2) < 1$, 此级数收敛.

定理 2 $q_J(a)$ 关于 $s(a)$ 是凸的, 因而当然是连续的.

证明 利用强 Markoff 性, 当 $j_1 < a_1 < a < a_2 < j_2$ 时,

$$p_J(a) = p_{(a_1, a_2)}(a) + s(a; a_1, a_2) p_J(a_1) + s(a; a_2, a_1) p_J(a_2),$$

(这在 Dynkin 定理中已以更一般的情形证明). 以 $s(a)$ 表示 $s(a; a_1, a_2), s(a; a_2, a_1)$ 得

$$\begin{aligned} p_J(a) =& p_{(a_1, a_2)}(a) + \frac{s(a_2) - s(a)}{s(a_2) - s(a_1)} p_J(a_1) + \frac{s(a) - s(a_1)}{s(a_2) - s(a_1)} p_J(a_2) \\ \geqslant& \frac{s(a_2) - s(a)}{s(a_2) - s(a_1)} p_J(a_1) + \frac{s(a) - s(a_1)}{s(a_2) - s(a_1)} p_J(a_2). \end{aligned}$$

因此 $p_J(a)$ 是凹的, 而 $q_J(a) = -p_J(a)$ 是凸的.

由此定理可以决定 $m_J(a) = \mathrm{d}q_J(a)/\mathrm{d}s(a)$, 而且它是增函数.

定理 3 若 $J \subset J'$, 则 $m_J(a) \equiv m_{J'}(a)+$ 常数, $a \in J$.

证明 与定理 2 的证明一样, 对 $a \in J$,

$$p_{J'}(a) = p_J(a) + \frac{s(j_2) - s(a)}{s(j_2) - s(j_1)} p_J(j_1) + \frac{s(a) - s(j_1')}{s(j_2) - s(j_1)} p_J(j_2).$$

若关于 s 微分, 则第二项和第三项都为常数, 故

$$m_{J'}(a) = m_J(a) + \text{常数}.$$

§55 Feller 典范形

令 I 为 $x^{(a)}(t)$ 的正则区间, $s(x)$ 和 $dm(x)$ 分别为典范尺度和典范测度. 更令 A_I 为 $x^{(a)}(t)$ 在 I 上的局部生成算子. 证明 **Feller 典范形**:

$$A_I = (D_m D_s^+)_I \tag{55.1}$$

就是本节的目的.

先说明 $D_m D_s^+$ 的定义. 固定 x, 像普通的微分定义一样, 定义

$$D_s^+ f(x) = \lim_{s \downarrow 0} \frac{f(x+\varepsilon) - f(x)}{s(x+\varepsilon) - s(x)},$$

$$D_m D_s^+ f(x) = \lim_{\varepsilon, \varepsilon' \downarrow 0} \frac{D_s^+ f(x+\varepsilon) - D_s^+ f(x-\varepsilon')}{m(x+\varepsilon) - m(x-\varepsilon')}.$$

要下此定义, 只需在 x 的邻域定义 f 即可. 令 f 为定义于开区间 I 的函数, 如对所有的 $x \in I$, 可确定上面的 $D_m D_s^+ f(x)$, 且对 x 连续, 则定义

$$f \in \mathfrak{D}((D_m D_s^+)_I), \quad (D_m D_s^+)_I f(x) = D_m D_s^+ f(x). \tag{55.2}$$

因为, f 当然必须是连续的. (55.1) 肯定, 两边的算子的定义域与值域是一致的.

引理 1 $s \in \mathfrak{D}(A_I)$, 且 $A_I s(x) = 0$.

证明 因 0 可以看作连续函数, 故只要对任意 b 证明 $A_b s = 0$ 即可. 采用 §51 定义 2 定义 A_b. 令 b 的邻域 $J = (j_1, j_2)$ 为使 $\overline{J} \subset I$ 的区间. 若令 $s_J(x) = s(x; j_2, j_1)$, 则可以写成 $s(x) = \alpha s_J(x) + \beta$, 故只要证明 $A_b s_J = 0$. 由 Markoff 性,

$$\begin{aligned} s_J(b) &= \int_J P_J(t, b, \mathrm{d}y) s(y) + P(\tau_J^{(b)} \leqslant t, x^{(b)}(\tau_J^{(b)}) = j_2) \\ &= \int_J P_J(t, b, \mathrm{d}y) s_J(y) + o(t). \end{aligned}$$

故

$$\frac{1}{t}\left\{ \int_J P_J(t, b, \mathrm{d}y) s_J(y) - s_J(b) \right\} = o(1),$$

即

$$A_b s_J = 0.$$

引理 2 令 $q(x) = \int_{x_0}^{x} m(y)\mathrm{d}s(y)(x_0$ 是 I 内任一定点, x 则为 I 内的动点), 则 $q \in \mathfrak{D}(A_r)$ 且 $A_I q(x) = 1$.

证明 与引理 1 一样, 对 $b \in I$ 证明 $A_b q = 1$ 即可. 考虑 b 的邻域 J, $\overline{J} \subset I$, 则对 $x \in J$,

$$\frac{\mathrm{d}q_J(x)}{\mathrm{d}s(x)} = m(x) + 常数,$$

故 $q(x) = q_J(x) + \alpha s(x) + \beta$. 由引理 1, $A_b s = 0$, 故如能证 $A_b q_J = 1$, 即 $A_b p_J = -1$, 则本引理得证. 利用 Markoff 性,

$$\begin{aligned} p_J(b) &= \int_J P_J(t, b, \mathrm{d}y)(p_J(y) + t) + E(\tau_J^{(b)}; \tau_J^{(b)} < t) \\ &= \int_J P_J(t, b, \mathrm{d}y)p_J(y) + P_J(t, b, J) \cdot t + t \cdot o(t) \\ &= \int_J P_J(t, b, \mathrm{d}y)p_J(y) + (1 - o(t)) \cdot t + t \cdot o(t). \end{aligned}$$

$$\frac{1}{t}\left\{ \int_J P_J(t, b, \mathrm{d}y)p_J(y) - p_J(b) \right\} = -1 + o(t) \to -1.$$

引理 3 若 $f \in \mathfrak{D}(A_I)$, 且 $A_I f(b) > 0$, 故在 b 的某邻域内, $f(x)$ 关于 $s(x)$ 是凸的.

证明 因为 $A_I f$ 连续而且 $A_I f(b) > 0$, 故在 b 的某邻域 J 上 $A_I f(x) > 0$. 在 J 内任取两点 a_1, a_2, 又定出 α, β 使

$$f(a_i) = \alpha s(a_i) + \beta, \quad i = 1, 2,$$

则只要证明, 对 $a_1 \leqslant x \leqslant a_2$,

$$f(x) \leqslant \alpha s(x) + \beta$$

即可. 因为 $g(x) \equiv f(x) - \alpha s(x) - \beta$ 是连续函数, 故在 $a_1 \leqslant x \leqslant a_2$ 上取最大值, 设为 $g(a_0)$. 当 $a_0 = a_1$ 或 $a_0 a_2$ 时, 上式 (化为等式) 成立. 若 $a_0 \in (a_1, a_2)$, 则因 $g(x)$ 在 a_0 取极大值, 故 $A_I g(a_0) \leqslant 0$, 从而

$$A_I f(a_0) \leqslant \alpha A_I s(a_0) + \beta A_I \cdot 1 = 0.$$

此与 $A_I f(x) > 0 (x \in J)$ 矛盾.

定理 1 若 $f \in \mathfrak{D}(A_I)$, 则 $f \in \mathfrak{D}((D_m D_s^+)_I)$, 且在 I 内,

$$A_I f = (D_m D_s^+)_I f,\ \text{即} A_x f = D_m D_s^+ f(x), x \in I.$$

证明 令 $f \in \mathfrak{D}(A_I), \alpha = A_I f(b)$, 并令

$$g(x) = f(x) - (\alpha - \delta) q(x),$$

则

$$A_I g(b) = \alpha - (\alpha - \delta) \cdot 1 = \delta > 0,$$

故在 b 的某邻域 J 上 g 关于 s 是凸的, 因而 $D_s^+ g(x)$ 在 J 内是增函数. 故若在 b 的两边取 J 的点 $b_1, b_2 (b_2 > b > b_1)$, 则

$$D_s^+ g(b_2) > D_s^+ g(b_1),$$

因此

$$D_s^+ f(b_2) - D_s^+ f(b_1) > (\alpha - \delta)(m(b_2) - m(b_1)).$$

故得

$$\frac{D_s^+ f(b_2) - D_s^+ f(b_1)}{m(b_2) - m(b_1)} > \alpha - \delta.$$

由此得

$$\lim_{\varepsilon, \varepsilon' \downarrow 0} \frac{D_s^+ f(b + \varepsilon) - D_s^+ f(b - \varepsilon')}{m(b + \varepsilon) - m(b - \varepsilon')} \geqslant \alpha.$$

同样取 $\alpha + \delta$ 代替 $\alpha - \delta$ 就可证 $\lim_{s, s' \downarrow 0} \leqslant \alpha$. 因此

$$D_m D_s^+ f(b) = A_I f(b).$$

因为右边关于 b 连续, 故得 $(D_m D_s^+)_I f(b) = A_I f(b)$.

定理 1 的逆定理如下.

定理 2 若 $f \in \mathfrak{D}((D_m D_s^+)_I)$, 则 $f \in \mathfrak{D}(A_I)$, 且在 I 内,

$$(D_m D_s^+)_I f = A_I f.$$

证明 设 R 的点全部都是 $x^{(a)}(t)$ 的一维扩散点, 且设 I 是 $x^{(a)}(t)$ 的正则区间. 首先定义 L_b 如下:

$$b \in I \to L_b f = D_m D_s^+ f(b), \quad b \notin I \to L_b f = A_b f.$$

其次定义 L 为

$$\mathfrak{D}(L) = \{f/\text{对所有}b, f \in \mathfrak{D}(L_b), \text{而且}L_b f\text{对}b \in R\text{连续}\},$$
$$Lf(b) \equiv L_b f.$$

若 $f \in \mathfrak{D}(A)$, 则 $f \in \mathfrak{D}(A_I)$, 故由定理 1, $f \in \mathfrak{D}((D_m D_s^+)_I)$, 若 $b \in I$, 则 $f \in \mathfrak{D}(L_b)$, 而且

$$L_b f = D_m D_s^+ f(b) = A_b f = Af(b).$$

若 $b \notin I$, 则由 $f \in \mathfrak{D}(A)$ 得 $f \in \mathfrak{D}(A_b)$, 故 $f \in \mathfrak{D}(L_b)$, 而且 $L_b f = A_b f = Af(b)$. 因为 $Af(b)$ 关于 b 连续, 故 $f \in \mathfrak{D}(L)$, 而且 $Lf = Af$, 故 $L > A$.

其次证 $(\lambda - L)^{-1}(\lambda > 0)$ 的存在. 为此, 只要从 $\lambda u = Lu$ 得出 $u = 0$ 即可. 令 u 的最小值为 $u(a)$, 试证 $u(a) \geqslant 0$. 若 $a \in I$, 则在 a 的邻域上 $Lu(x) = D_m D_s^+ u(x)$. 若 $u(a) < 0$, 则 $Lu(a) < 0$, 故在 a 的邻域上 $D_m D_s^+ u(x) < 0$, 故 $D_s^+ u(x)$ 下降, $-u(x)$ 关于 $s(x)$ 为凸. 因此 $-u(x)$ 不能于 a 取最大值, 即 $u(x)$ 不能于 a 取最小值, 这与假设矛盾, 即有 $u(a) \geqslant 0$. 若 $a \notin I$, 则 $Lu(a) = A_a u$ 而且 $u(a)$ 最小, 故由 A_a 的定义得 $A_a u \geqslant 0$, 因此 $u(a) = \lambda^{-1} Lu(a) \geqslant 0$. 于是无论如何 $u(a) \geqslant 0$, 即 $u \geqslant 0$. 又如以 $-u$ 代 u 就得 $u \leqslant 0$, 故 $u = 0$.

由 $L > A$ 得 $(\lambda - L)^{-1} > (\lambda - A)^{-1} = R_\lambda$. 因为 R_λ 的定义域的全体是 R, 故 $(\lambda - L)^{-1} = (\lambda - A)^{-1}$. 故 $L = A$.

令 $f \in \mathfrak{D}((D_m D_s^+)_I)$, b 为 I 的任意点, 若能确定 $A_b f$ 并且说明它关于 b 连续, 则定理证完. 能在 I 的任意子区间 $K = (k_1, k_2)(\overline{K} \subset I)$ 内证明 $A_b f$ 连续即可. 又可以令 K 的端点是 m 的连续点. 取满足 $\overline{K} \subset J \subset \overline{J} \subset I$ 的区间 $J = (j_1, j_2)$ 并令其端点也是 m 的连续点. 试构造 g 使其在 K 上与 f 相等, 在 J 外则为 0, 而且使 $D_m D_s^+ g(b)$ 关于 $b \in I$ 连续. 如能这样构造, 则 g 属于 L 即 A 的定义域, 因而 $A_b g = Ag(b)$ 连续于 $b \in J$. 因在 K 上 $f = g$, 故利用局部性得 $A_b f = A_b g$, 于是 $A_b f, b \in K$ 也连续. 剩下来的就是 g 的构造. 也可以构造 $h = Lg$ 来代替 g. 因在 K 上 $h = D_m D_s^+ f$, 在 J 外 $h = 0$, 故只要在 $[j_1, k_1], [k_2, j_2]$ 上定义它. 今说明在前一区间上如何构造, 由于"$D_s^+ h$ 在 m 的连续点上的连续性"和"$D_m D_s^+ h$ 的连续

性”, h 应满足的充要条件是:

$$\begin{aligned} l_1(h) &\equiv h(j_1) = 0, \\ l_2(h) &\equiv h(k_1) = D_m D_s^+ f(k_1), \\ l_3(h) &\equiv \int_{j_1}^{k_1} h(x)\mathrm{d}m(x) = D_s^+ f(k_1), \\ l_4(h) &\equiv \int_{j_1}^{k_1} \int_{j_1}^{y} h(x)\mathrm{d}m(x)\mathrm{d}s(y) \\ &\equiv \int_{j_1}^{k_1} h(x)(s(k_1) - s(x))\mathrm{d}m(x) = f(k_1). \end{aligned}$$

因为 $C[j_1, k_1]$ 上的泛函 l_1, l_2, l_3, l_4 显然线性无关, 故这样的 h 的确存在. 在 $[k_2, j_2]$ 上也是如此. 于是定理 2 完全得证.

§56　一般通过点上的局部生成算子

R 的开集 I, 若能与实数的开集 (r_1, r_2) 同胚, 则简称为 R 的开区间. a 称为开区间 I 的端点, 就是指 $\{a\} \cup I$ 与半开区间 $[r_1, r_2)$(或 $(r_1, r_2]$) 同胚. 在这个拓扑对应中, a 当然变换到 r_1(或 r_2).

令 a 为只由扩散点构成的开区间 I 的端点. 对 a 的充分小的邻域 U, 如有

$$P(\text{对}0 \leqslant t < \tau_U^{(a)}\text{中之}t, x^{(a)}(t) \in \{a\} \cup I) = 1, \tag{56.1}$$

则称a为**一般通过点**.一般通过点仅限于它是一维点时才成为通过点 (§52).

在通过点的左右分类中, 点的邻域同胚映射于实数区间的映象方式与方向有关. 同样, 对一般通过点, 若把 $\{a\} \cup I$ 映射于 $[r_1, r_2)$, 就是**一般右通过点**, 而映射于 $(r_1, r_2]$ 时, 就是**一般左通过点**. 以后仅讨论一般右通过点.

一般右通过点仍有可能是套点, 若把这种过于简单的情况除外, 取 a 的充分小的邻域 U, 则除 (56.1) 外还可假定

$$E(\tau_U^{(b)}) < \infty, \quad b \in U. \tag{56.2}$$

设把 $\{a\} \cup I$ 内的点与所对应的 $[r_1, r_2)$ 内的点看成同一点. 因接近 r_1 的点位于 U 内, 故使 $[r_1, \xi)$ 包含于 U 内的 ξ 的上确界 $r_2' > r_1$.

若对 $r_1 < \xi < r_2'$, 令

$$p(\xi) = E(\tau_{[r_1,\xi)}^{(r_1)}), \tag{56.3}$$

则 $0 \leqslant p(\xi) < \infty$. 因为 $x^{(r_1)}(t)$ 当 $0 \leqslant t < \tau_U^{(r_1)}$ 时常处于 $[r_1, r_2)$ 内, 而且 t 在 $\tau_U^{(r_1)}$ 的附近时接近于 r_2', 故存在这样的 $t : x^{(r_1)}(t) = \xi, 0 \leqslant t < \tau_U^{(r_1)}$. 令满足这种关系的最小的 t 为 τ^*, 对它用强 Markoff 性, 则得

$$p(\xi) + E(\tau_U^{(\xi)}) = E(\tau_U^{(r_1)}). \tag{56.4}$$

因为 $[r_1, r_2')$ 是一维扩散点的集合, 故与正则点的情况 (§55) 一样, 可以证明 $p_U(\xi) = E(\tau_U^{(\xi)})$ 满足

$$A_\xi p_U = -1.$$

故得下述结论.

引理 1

$$A_\xi p = 1. \tag{56.5}$$

引理 2 若 $f \in \mathfrak{D}(A_v)$, V 是 a 的任意邻域, 且

$$A_V f(r_1) > 0,$$

则对足够小的 $\varepsilon(>0)$,

$$f(\xi) \geqslant f(r_1), \quad r_1 < \xi < r_1 + \varepsilon.$$

证明 因为 $A_V f(r_1) > 0$, 故对足够小的 ε,

$$A_V f(\xi) > 0, \quad r_1 < \xi < r_1 + \varepsilon.$$

因此 $f(\xi)$ 不能在 $r_1 < \xi < r_1 + \varepsilon$ 内取极大. 故 $f(\xi)$ 在 $\xi = r_1$ 的某邻域 $(r_1 \leqslant \xi < r_1 + \delta)$ 上或者非增, 或者非减. 如是前一情况, 则 $A_{r_1} f = 0$, 即 $A_V f(r_1) = 0$, 此与假设矛盾.

定理 1 若 $f \in \mathfrak{D}(A_V)$, 则

$$A_V f(r_1) = \lim_{\xi \downarrow r_1} \frac{f(\xi) - f(r_1)}{p(\xi)}.$$

证明 若令 $A_V f(r_1)=\alpha$, 则 $g(\xi)=f(\xi)-(\alpha-\varepsilon)p(\xi), r_1<\xi<r_2'$, 满足 $g\in\mathfrak{D}(A_V)$, 并且

$$A_V g(r_1)=\alpha-(\alpha-\varepsilon)\cdot 1=\varepsilon>0.$$

故 g 在 r_1 的右边增加, 因而

$$g(\xi)\geqslant g(r_1),\quad r_1<\xi<r_1+\delta,$$

于是

$$\frac{f(\xi)-f(r_1)}{p(\xi)}>\alpha-\varepsilon,\quad r_1<\xi<r_1+\delta,$$

故

$$\varliminf_{\xi\downarrow r_1}\frac{f(\xi)-f(r_1)}{p(\xi)}\geqslant\alpha.$$

同样

$$\varlimsup_{\xi\downarrow r_1}\frac{f(\xi)-f(r_1)}{p(\xi)}\leqslant\alpha.$$

§57 最初通过时间的分布

令 $I=(r_1,r_2)$ 为 $x^{(a)}(t)$ 的正则区间, s 和 $\mathrm{d}m$ 分别为其典范尺度与典范测度. 取 I 的子区间 $J=(j_1,j_2),\overline{J}\subset J$, 且令 a 为 J 内的任意点. 令 $\tau_J^{(a)}$ 为 $x^{(a)}(t)$ 在 J 的最初通过时间. 易见 $\tau_J^{(a)}$ 的平均值有限, 因而它以概率 1 为有限的. $x^{(a)}(\tau_J^{(a)})$ 就是 j_1 或 j_2, 根据这两种情况分别规定

$$\tau_1^{(a)}=\tau_J^{(a)},\quad \tau_2^{(a)}=\infty,$$

或

$$\tau_2^{(a)}=\tau_J^{(a)},\quad \tau_1^{(a)}=\infty.$$

$\tau_i^{(a)}<\infty$ 等价于 $\tau_i^{(a)}=\tau_J^{(a)}$, 即 $x^{(a)}(\tau_J^{(a)})=j_i$. 如 §55 所证,

$$A_J=(D_m D_s^+)_J. \tag{57.1}$$

$s_J(a)=s(a;j_2,j_1)=P(\tau_2^{(a)}<\infty)$, 而且

$$\left.\begin{array}{l} A_J s_J(a)=0,\\ s_J(j_1+0)=0,\quad s_J(j_2-0)=1. \end{array}\right\} \tag{57.2}$$

这由 (57.1) 及 $s_J(a)=(s(a)-s(j_1))/(s(j_2)-s(j_1))$ 易得.

其次, 若令 $p_J(a)=E(\tau_J^{(a)})$, 则如 §55 所述,

$$A_J p_J(a)=-1, \tag{57.3}$$

进而还可证明

$$p_J(j_1+0)=p_J(j_2-0)=0. \tag{57.3$'$}$$

实际上, 因 j_1 是正则点, 故对其充分小的邻域 U, 得 $p_U(x)<\varepsilon, x\in U$. $p_U(x)=E(\tau_U^{(x)})$. 取一充分小 δ, 使 $K=[j_1,j_1+\delta]\subset U$. 对 $x\in K$ 有

$$p_J(x)=p_K(x)+s(x;\xi,j_1)p_J(\xi)<\varepsilon+s(x;\xi,j_1)p_J(\xi).$$

若令 $x\to j_1$, 则 $s(x;\xi,j_1)\to 0$. 于是

$$p_J(j_1+0)\leqslant\varepsilon,\ \text{故}p_J(j_1+0)=0.$$

同样

$$p_J(j_2-0)=0.$$

定理 1 s_J 及 p_J 分别由 (57.2) 和 (57.3), (57.3′) 决定.

证明 由上述易知 s_J 及 p_J 分别满足这些条件. 又由 (57.1) 知 $A_Ju=0$ 有两个线性独立的解, 因而由上述的边界条件, 可知解是唯一决定的.

以后对 $i=1,2$, 令

$$\varphi_i(\mathrm{d}t,a)=P(\tau_i^{(a)}\in\mathrm{d}t),$$

$$\widehat{\varphi}_{i\lambda}(a)=\int_0^\infty \mathrm{e}^{-\lambda t}\varphi_i(\mathrm{d}t,a)=E(\mathrm{e}^{-\lambda\tau_i^{(a)}}).$$

定理 2 $\widehat{\varphi}_{1\lambda}$ 是

$$A_J\widehat{\varphi}_{1\lambda}=\lambda\widehat{\varphi}_{1\lambda},\quad \widehat{\varphi}_{1\lambda}(j_1+0)=1,\quad \widehat{\varphi}_{1\lambda}(j_2-0)=0 \tag{57.4}$$

的唯一解; 同样 $\widehat{\varphi}_{2\lambda}$ 是

$$A_J\widehat{\varphi}_{2\lambda}=\lambda\widehat{\varphi}_{2\lambda},\quad \widehat{\varphi}_{2\lambda}(j_2-0)=1,\quad \widehat{\varphi}_{2\lambda}(j_1+0)=0 \tag{57.5}$$

的唯一解.

证明 考虑 $\widehat{\varphi}_{1\lambda}$. 如前, 令

$$P_J(t,a,\mathrm{d}y)=P_J(x^{(a)}(t)\in\mathrm{d}y,\tau_J^{(a)}>t).$$

$$\varphi_1(\mathrm{d}t+s,a)=\int_J P_J(s,a,\mathrm{d}y)\varphi_1(\mathrm{d}t,y),$$

$$\int_0^\infty \mathrm{e}^{-\lambda t}\varphi_1(\mathrm{d}t+s,a)=\int_J P_J(s,a,\mathrm{d}y)\int_0^\infty \mathrm{e}^{-\lambda t}\varphi_1(\mathrm{d}t,y),$$

$$\mathrm{e}^{\lambda s}\int_s^\infty \mathrm{e}^{-\lambda t}\varphi_1(\mathrm{d}t,a)=\int_J P_J(s,a,\mathrm{d}y)\widehat{\varphi}_{1\lambda}(y).$$

但

$$\int_s^\infty \mathrm{e}^{-\lambda t}\varphi_1(\mathrm{d}t,a)=\widehat{\varphi}_{1\lambda}(a)-\int_0^s \mathrm{e}^{-\lambda t}\varphi_1(\mathrm{d}t,a),$$

$$\int_0^s \mathrm{e}^{-\lambda t}\varphi_1(\mathrm{d}t,a)\leqslant \varphi_1([0,s],a)\leqslant P(\tau_U^{(a)}<s)=o(s).$$

故

$$\mathrm{e}^{\lambda s}(\widehat{\varphi}_{1\lambda}(a)+o(s))=\int_J P_J(s,a,\mathrm{d}y)\widehat{\varphi}_{1\lambda}(y).$$

由此立得

$$A_a\widehat{\varphi}_{1\lambda}=\lambda\widehat{\varphi}_{1\lambda}(a).$$

如能证明右边关于 a 是连续的, 则即已证明在 J 上

$$A_J\widehat{\varphi}_{1\lambda}=\lambda\widehat{\varphi}_{1\lambda}.$$

现在, 像 $\widehat{\varphi}_{1\lambda}(b)$ 定义于 J 上一样, 若定义于另一区间 K 上时, 就记为 $\widehat{\varphi}_{1\lambda}(b;K)$. 依照以前多次用到的强 Markoff 性的证法, 当 $a<b$ 时,

$$\widehat{\varphi}_{1\lambda}(b)=\widehat{\varphi}_{1\lambda}(b;(a,j_2))\widehat{\varphi}_{1\lambda}(a)\leqslant \widehat{\varphi}_{1\lambda}(a).$$

故 $\widehat{\varphi}_{1\lambda}$ 单调下降. 当取 $a<c<b$, $c\in J$, 使 a,b 充分接近于 c 时, 若能证

$$\widehat{\varphi}_{1\lambda}(b;(a,j_2))>1-\varepsilon,$$

则得证 $\widehat{\varphi}_{1\lambda}$ 的连续性.

$$\begin{aligned}\widehat{\varphi}_{1\lambda}(b;(a,j_2))=&E\{\mathrm{e}^{-\lambda\tau^{(b)}_{(a,j_2)}};x^{(b)}(\tau^{(b)}_{(a,j_2)})=a\}\\ \geqslant&\mathrm{e}^{-\lambda t}P(\tau^{(b)}_{(a,j_2)}<t,x^{(b)}(\tau^{(b)}_{(a,j_2)})=a)\\ \geqslant&\mathrm{e}^{-\lambda t}\{P(\tau^{(b)}_{(a,k)}<t)-P(x^{(b)}(\tau^{(b)}_{(a,k)})=k)\}.\end{aligned}$$

其中 k 是 b 与 j_2 之间的点.

$$P(\tau_{(a,k)}^{(b)} \geqslant t) \leqslant \frac{1}{t} E(\tau_{(a,k)}^{(b)}),$$

$$P(x^{(b)}(\tau_{(a,k)}^{(b)}) = k) = \frac{s(b) - s(a)}{s(k) - s(a)}.$$

令 $t = \delta/\lambda$. 取 c 的充分小的邻域 U, 使 $E(\tau_U^{(b)}) < \delta t$, 在 U 内取 $a, b, k, a < c < b < k$, 使 a, b 极其接近, 以致 $(s(b) - s(a))/(s(k) - s(a)) < \delta$. 因为

$$E(\tau_{(a,k)}^{(b)}) \leqslant E(\tau_U^{(b)}) < \delta t,$$

故

$$\widehat{\varphi}_{1\lambda}(b; (a, j_2)) \geqslant \mathrm{e}^{-\lambda t}(1 - 2\delta),$$

取 t 足够小, 则上式右边

$$> (1 - \delta)(1 - 2\delta) > 1 - \varepsilon(\delta = \frac{\varepsilon}{3}).$$

故得证 $\widehat{\varphi}_{1\lambda}(a)$ 的连续性.

剩下来的是要证 $\widehat{\varphi}_{1\lambda}(j_1 + 0) = 1, \widehat{\varphi}_{1\lambda}(j_2 - 0) = 0$ 以及唯一性. $\widehat{\varphi}_{1\lambda}(a)$ 随 a 增加而减少, 并且处于 0 与 1 之间. 若在 j_1 的左边取 k_1, 在 a 与 j_2 之间取 k_2, 则像证连续性一样, 可得

$$\widehat{\varphi}_{1\lambda}(a) \geqslant \mathrm{e}^{-\lambda t}\{P(\tau_{(k_1,k_2)}^{(a)} < t) - P(x^{(a)}(\tau_{(k_1,k_2)}^{(a)}) = k_2)\}.$$

由此可见当 a 充分接近于 j_1 时, $\widehat{\varphi}_{1\lambda}(a) > 1 - \varepsilon$, 故 $\widehat{\varphi}_{1\lambda}(j_1 + 0) = 1$. 同样 $\widehat{\varphi}_{2\lambda}(j_2 - 0) = 1$. 由于

$$\widehat{\varphi}_{1\lambda}(a) + \widehat{\varphi}_{2\lambda}(a) = E(\mathrm{e}^{-\lambda \tau_J^{(a)}}) \leqslant 1,$$

故

$$\widehat{\varphi}_{1\lambda}(j_2 - 0) \leqslant 1 - \widehat{\varphi}_{2\lambda}(j_2 - 0) = 0, \quad 即 \widehat{\varphi}_{1\lambda}(j_2 - 0) = 0.$$

为证唯一性, 只要由 $A_J u = \lambda u, u(j_1 + 0) = u(j_2 - 0) = 0$ 推出 $u \equiv 0$ 即可. 如在某些点上 $u > 0$, 则 u 在 J 内取正的最大值, 设为 $u(a)$, 则 $A_J u(a) \leqslant 0$, 与 $u(a) \leqslant 0$ 矛盾, 故 $u \leqslant 0$. 同样证明 $u \geqslant 0$, 因此 $u \equiv 0$.

定理 3 令 $\widehat{\varphi}_\lambda(a) = E(\mathrm{e}^{-\lambda \tau_J^{(a)}})$, 则它是

$$A_J \widehat{\varphi}_\lambda = \lambda \widehat{\varphi}_\lambda, \quad \widehat{\varphi}_\lambda(j_1 + 0) = \widehat{\varphi}_\lambda(j_2 - 0) = 1$$

的唯一解.

证明 由定理 2 所用的公式

$$\widehat{\varphi}_\lambda(a) = E(\mathrm{e}^{-\lambda\tau_J^{(a)}}) = \widehat{\varphi}_{1\lambda}(a) + \widehat{\varphi}_{2\lambda}(a),$$

即得

$$A_J\widehat{\varphi}_\lambda = A_J\widehat{\varphi}_{1\lambda} + A_J\widehat{\varphi}_{2\lambda} = \lambda\widehat{\varphi}_{1\lambda} + \lambda\widehat{\varphi}_{2\lambda} = \lambda\widehat{\varphi}_\lambda.$$

又

$$\widehat{\varphi}_\lambda(j_1+0) = \widehat{\varphi}_{1\lambda}(j_1+0) + \widehat{\varphi}_{2\lambda}(j_1+0) = 1,$$

$$\widehat{\varphi}_\lambda(j_2-0) = \widehat{\varphi}_{1\lambda}(j_2-0) + \widehat{\varphi}_{2\lambda}(j_2-0) = 1.$$

故 $\widehat{\varphi}_\lambda$ 满足定理中的方程及边界条件. 解的唯一性证明与定理 2 一样.

§58 古典扩散过程

令 $\mathbb{R}$ 为实数集合 $\mathbb{R}^1$ 或其区间 $[r_1, r_2]$. 令 $x^{(a)}(t)$ 为 $\mathbb{R}$ 上的扩散过程. 由 Ray 定理, 有

$$P\{|x^{(\xi)}(t)-\xi| > \varepsilon\}/t \to 0 \quad (t \to 0). \tag{58.1}$$

若假定

$$a(\xi) = \lim_{t\downarrow 0} \frac{E\{x^{(\xi)}(t)-\xi; |x^{(\xi)}(t)-\xi| < \varepsilon\}}{t}, \tag{58.2}$$

$$b(\xi) = \lim_{t\downarrow 0} \frac{E\{(x^{(\xi)}(t)-\xi)^2; |x^{(\xi)}(t)-\xi| < \varepsilon\}}{t} \tag{58.3}$$

对某 $\varepsilon > 0$ 存在, 则由 (58.1), 对所有的 $\varepsilon > 0$ 都存在, 而且其值与 ε 无关. 若 $a(\xi)$ 与 $b(\xi)$ 都是 $\xi \in (r_1, r_2)$ 的连续函数, 则 $x^{(a)}(t)$ 叫做 $\mathbb{R}(=(r_1, r_2])$ 上的古典扩散过程或 Kolmogoroff 扩散过程. 利用 $x^{(a)}(t)$ 的转移概率 $P(t, a, E)$, 可把 $a(\xi), b(\xi)$ 写为

$$a(\xi) = \lim_{t\downarrow 0} \frac{1}{t} \int_{|y-\xi|<\varepsilon} (y-\xi) P(t, \xi, \mathrm{d}y), \tag{58.2$'$}$$

$$b(\xi) = \lim_{t\downarrow 0} \frac{1}{t} \int_{|y-\xi|<\varepsilon} (y-\xi)^2 P(t, \xi, \mathrm{d}y). \tag{58.3$'$}$$

$a(\xi)$ 可取正负值, 但 $b(\xi) \geqslant 0$.

例如, 对 Wiener 过程, 求得 $a(\xi), b(\xi)$ 为

$$a(\xi) = 0, \quad b(\xi) = 1. \tag{58.4}$$

定理 1 若 f 是 $J = (j_1, j_2)(r_1 \leqslant j_1 < j_2 \leqslant r_2)$ 内二次连续可微函数, 则 $f \in \mathfrak{D}(A_J)$, 而且

$$A_J f(\xi) = (a(\xi)\frac{\mathrm{d}}{\mathrm{d}\xi} + \frac{b(\xi)}{2}\frac{\mathrm{d}^2}{\mathrm{d}\xi^2})f(\xi). \tag{58.5}$$

证明 因上式右边是 ξ 的连续函数, 故只要证明 $A_\xi f$ 等于右边即可. 对 $\xi \in J, \delta > 0$, 取 $\varepsilon >$ 充分小, 就 $|y - \xi| < \varepsilon$ 而论, 有

$$f(y) - f(\xi) = f'(\xi)(y - \xi) + \frac{f''(\xi) \pm \delta}{2}(y - \xi)^2,$$

$$\begin{aligned}
&\int_{|y-\xi|<\varepsilon} f(y)P(t, \xi, \mathrm{d}y) - f(\xi) \\
=&\int_{|y-\xi|<\varepsilon} (f(y) - f(\xi))P(t, \xi, \mathrm{d}y) + o(t) \\
=&f'(\xi)\int_{|y-\xi|<\varepsilon} (y - \xi)P(t, \xi, \mathrm{d}y) \\
&+ \frac{f''(\xi) \pm \delta}{2}\int_{|y-\xi|<\varepsilon} (y - \xi)^2 P(t, \xi, \mathrm{d}y) + o(t) \\
=&f'(\xi)a(\xi) \cdot t + \frac{f''(\xi) \pm \delta}{2} b(\xi) \cdot t + o(t).
\end{aligned}$$

故得

$$\overline{\lim}\frac{1}{t}\left\{\int_{|y-\xi|<\varepsilon} f(y)P(t, \xi, \mathrm{d}y) - f(\xi)\right\} \leqslant a(\xi)f'(\xi) + \frac{b(\xi)}{2}f''(\xi) + \frac{b(\xi)}{2} \cdot \delta.$$

由 (58.1), 左边与 ε 无关, 故右边的 δ 可以任意小, 因此这一极限 $\leqslant af' + \frac{b}{2}f''$. 同样取下极限就有 $\geqslant af' + \frac{b}{2}f''$. 结果得 $f \in \mathfrak{D}(A_\xi)$, 而且 $A_\xi f = af'(\xi) + \frac{b}{2}f''(\xi)$.

定理 2 若 $b(\xi) > 0$, 则 ξ 是正则点.

证明 若 ξ 是右通过点, 则对 $\mathfrak{D}(A_\xi)$ 中在 ξ 的邻域 J 上增加的函数 f 有 $A_\xi f \geqslant 0$. 但若在 ξ 的邻域上定义 f 为

$$f(x) = (x - \xi) - \beta(x - \xi)^2,$$

则

$$\left(af'+\frac{b}{2}f''\right)(\xi)=a-\beta b.$$

若令 $\beta>a(\xi)/b(\xi)$, 则上式为负, 故得

$$A_\xi f=A_J f(\xi)=\left(af'+\frac{b}{2}f''\right)(\xi)<0,$$

而发生矛盾. 因此不能是右通过点. 同样证明它不能是左通过点. 故是正则点.

现在, 对于在 J 上二次连续可微的函数, 定义算子 D_J 为

$$D_J f(\xi)=a(\xi)f'(\xi)+\frac{b(\xi)}{2}f''(\xi). \tag{58.6}$$

定理 3　若在 J 上 $b(\xi)>0$, 则

$$A_J=D_J. \tag{58.7}$$

证明　由定理 1, $A_J>D_J$. 将 D_J 变形得

$$\begin{aligned}D_J f=&\frac{b}{2}\mathrm{e}^{-B}\left(\mathrm{e}^B\frac{2a}{b}f'+\mathrm{e}^Bf''\right)\quad\left(B=\int\frac{2a}{b}\mathrm{d}\xi\right)\\=&\frac{b}{2}\mathrm{e}^{-B}(\mathrm{e}^Bf')'=\frac{1}{\frac{2}{b}\mathrm{e}^B}\frac{\mathrm{d}}{\mathrm{d}\xi}\frac{1}{\mathrm{e}^{-B}}\frac{\mathrm{d}}{\mathrm{d}\xi}f(\xi)\\=&\frac{\mathrm{d}}{\mathrm{d}m_1}\frac{\mathrm{d}}{\mathrm{d}s_1}f(\xi)\quad\left(s_1=\int\mathrm{e}^{-B}\mathrm{d}\xi,\quad m_1=\int\frac{2}{b}\mathrm{e}^B\mathrm{d}\xi\right).\end{aligned}$$

因为 $A_J=(D_mD_s^+)_J$, 故 $A_Ju=0$ 的全体解为 $\{\alpha+\beta s\}$, $D_Ju=0$ 的全体解为 $\{\alpha+\beta s_1\}$. 由于 $A_J>D_J$, 故前者包含后者. 因为二者都是二维的线性空间, 所以一致. 故得 $s=\alpha+\beta s_1$. 其次, 若令

$$q=\int m_1\mathrm{d}s_1=\int m_1\mathrm{e}^{-B}\mathrm{d}\xi,$$

则 $D_Jq=1$, 故 $A_Jq=1$, 即 $D_mD_s^+q=1$.

$$q=\int m\mathrm{d}s+\gamma s+\delta=\beta\cdot\int m\mathrm{d}s_1+\gamma's_1+\delta',$$

$$m_1=\frac{\mathrm{d}q}{\mathrm{d}s_1}=\beta m+\gamma',\quad m=\gamma''+\frac{1}{\beta}m_1.$$

故可以考虑 $m=m_1, s=s_1$. 若 $f\in\mathfrak{D}(A_J)$, 则 $D_{m_1}D_{s_1}^+f(\xi)$ 连续. 因为 $\frac{2}{b}\cdot\mathrm{e}^B$ 是连续而且是正的, 故上述连续性实际上就表示 $(\mathrm{e}^{-B}f')'$ 的连续性. 故 e^Bf' 是连续可微的. 由于 $(\mathrm{e}^{-B})'=-\mathrm{e}^{-B}\cdot\frac{2a}{b}$, 故 e^{-B} 是连续可微的, 从而 f' 也如此, 即 f 是二次连续可微的. 故 $f\in\mathfrak{D}(D_J)$, 于是 A_J 与 D_J 连同定义域一起是一致的.

例 对 Wiener 过程,

$$D_J=\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2},\quad \mathfrak{D}(D_J)=C_2(J).$$

由上述定理, 此算子即 A_J. 令 $J=(j_1,j_2)$ 为有限区间; 且令 $s_J(x)$ 为自 x 出发先到达右端点的概率, 则由上所述, $s_J(x)$ 是

$$A_Ju=0,\quad u(j_2)=1,\quad u(j_1)=0$$

的解. 由于 $A_J=D_J$, 故 $A_Ju=0$ 表示 $u''=0$, 即 $u=\alpha x+\beta$. 故得

$$s_J(x)=\frac{x-j_1}{j_2-j_1}.$$

这表示 x 就是典范尺度.

又 $P_J(x)=E(\tau_J^{(x)})$ 是

$$A_Ju=-1,\quad u(j_1)=u(j_2)=0$$

的解. 由 $A_Ju=-1$, 即 $u''=-2$, 得 $u=-x^2+cx+d$, 由 $u(j_1)=u(j_2)=0$ 定出 c,d, 得

$$u=(j_2-x)(x-j_1).$$

典范测度 $\mathrm{d}m$ 就是

$$m=-\frac{\mathrm{d}p_J}{\mathrm{d}s}=-\frac{\mathrm{d}p_J}{\mathrm{d}x}=2x,\quad 故\mathrm{d}m=2\mathrm{d}x.$$

若要求上节中定义的 $\hat{\varphi}_\lambda(x)=E(\mathrm{e}^{-\lambda\tau_J^{(x)}})$, 只要解

$$\frac{1}{2}\lambda u''=\lambda u,\quad u(j_1)=u(j_2)=1,$$

得

$$u=\alpha\mathrm{e}^{\sqrt{2\lambda}x}+\beta\mathrm{e}^{-\sqrt{2\lambda}x},$$

而 α,β 由 $u(j_1)=u(j_2)=1$ 决定.

§59 关于 Feller 算子 $D_mD_s^+$ 的端点的分类

暂时放下 Markoff 过程的概念, 一般地假定在区间 $I=(r_1,r_2)$ 上已定义连续、狭义增函数 $s(x)$ 与右连续增函数 m. 像以前一样, 定义 $(D_mD_s^+)_I$, 并称之为 **Feller 算子**. 现在, 根据 $D_mD_s^+$ 把 I 的端点 r_1,r_2 分为四种: **正则边界**(regular boundary)、**流出边界**(exit boundary)、**流入边界**(entrance boundary) 和**自然边界**(natural boundary). 为此首先引入下面的数:

$$\sigma_1=\iint\limits_{r_1<y<x<r_1'}\mathrm{d}m(x)\mathrm{d}s(y),\quad \mu_1=\iint\limits_{r_1<y<x<r_1'}\mathrm{d}s(x)\mathrm{d}m(y),$$

$$\sigma_2=\iint\limits_{r_2>y>x>r_2'}\mathrm{d}m(x)\mathrm{d}s(y),\quad \mu_2=\iint\limits_{r_2>y>x>r_2'}\mathrm{d}s(x)\mathrm{d}m(y).$$

当

$$\sigma_i<\infty,\quad \mu_i<\infty\text{时, 称}r_i\text{为正则边界},$$

$$\sigma_i<\infty,\quad \mu_i=\infty\text{时, 称}r_i\text{为流出边界},$$

$$\sigma_i=\infty,\quad \mu_i<\infty\text{时, 称}r_i\text{为流入边界},$$

$$\sigma_i=\infty,\quad \mu_i=\infty\text{时, 称}r_i\text{为自然边界}.$$

σ_i,μ_i 虽与 r_i' 的选择有关, 但它们是有限的或无限的却与此选择无关, 因而上述的分类也是这样.

这种分类的概率意义留在以后讲, 现在举出四种边界的例子.

例 1 设 $D_mD_s^+=\frac{\mathrm{d}^2}{\mathrm{d}x^2}$, $I=(-\infty,\infty)$, 则$m=x$, $s=x$, 故

$$\sigma_2=\iint\limits_{\infty>y>x>r_2'}\mathrm{d}x\,\mathrm{d}y=\int_{r_2'}^{\infty}(y-r_2')\mathrm{d}y=\infty,$$

$$\mu_2=\iint\limits_{\infty>y>x>r_2'}\mathrm{d}x\,\mathrm{d}y=\infty.$$

故 ∞ 是自然边界, 同样, $-\infty$ 也是自然边界.

例 2 若令 $D_mD_s^+=\frac{\mathrm{d}^2}{\mathrm{d}x^2}, I=(-\infty,0)$, 则 $-\infty$ 是自然边界, 0 是正则边界.

例 3 设 $D_mD_s^+ = \frac{\mathrm{d}}{\mathrm{d}x} + x^2\frac{\mathrm{d}^2}{\mathrm{d}x^2}, I = (0,2)$.

$$\frac{\mathrm{d}}{\mathrm{d}x} + x^2\frac{\mathrm{d}^2}{\mathrm{d}x^2} = \frac{\mathrm{d}}{x^{-2}\mathrm{e}^{-\frac{1}{x}}\mathrm{d}x}\frac{\mathrm{d}}{\mathrm{e}^{\frac{1}{x}}\mathrm{d}x},$$

故

$$\mathrm{d}m = x^{-2}\mathrm{e}^{-\frac{1}{x}}\mathrm{d}x,\ \mathrm{d}s = \mathrm{e}^{\frac{1}{x}}\mathrm{d}x.$$

因在 0 处有

$$\sigma = \iint\limits_{0<y<x<1} x^{-2}\mathrm{e}^{-\frac{1}{x}}\mathrm{e}^{\frac{1}{y}}\mathrm{d}x\ \mathrm{d}y = \infty,$$

$$\mu = \iint\limits_{0<y<x<1} y^{-2}\mathrm{e}^{-\frac{1}{y}}\mathrm{e}^{\frac{1}{x}}\mathrm{d}y\ \mathrm{d}x < \infty,$$

故 0 是流入边界. 易证 2 是正则的.

例 4 若 $D_mD_s^+ = -\frac{\mathrm{d}}{\mathrm{d}x} + x^2\frac{\mathrm{d}}{\mathrm{d}x^2}$, $I = (0,\infty)$, 则 0 是流出边界, ∞ 是自然边界.

§60 齐次方程 $(\lambda - D_mD_s^+)u = 0(\lambda > 0)$ 的特解

为方便计, 可设 $r_1 < 0 < r_2$, 也可以令

$$m(0-) = m(0+) = 0,\ s(0) = 0.$$

这个齐次方程的通解, 就是下述特解 e_0 和 e_1 的线性组合.

$$D_mD_s^+e_0 = \lambda e_0,\ e_0(0) = 1,\ D_s^+e_0(0) = 0, \tag{60.1}$$

$$D_mD_s^+e_1 = \lambda e_1,\ e_1(0) = 0,\ D_s^+e_1(0) = 1. \tag{60.2}$$

为求 e_0 与 e_1, 只需解下列积分方程

$$e_0(x) = 1 + \lambda\int_0^x\int_0^\xi e_0(\eta)\mathrm{d}m(\eta)\mathrm{d}s(\xi),$$

$$e_1(x) = s(x) + \lambda\int_0^x\int_0^\xi e_1(\eta)\mathrm{d}m(\eta)\mathrm{d}s(\xi).$$

因为 $x<0$ 的情形类似, 故只就 $x\geqslant 0$ 时来叙述求 $e_1(x)$ 的方法. 若令

$$p_1\circ p_2\circ\cdots\circ p_n(x)=\int\limits_{0\leqslant\xi_1\leqslant\cdots\leqslant\xi_n\leqslant x}\cdots\int \mathrm{d}p_1(\xi_1)\cdots\mathrm{d}p_n(\xi_n),$$

则

$$(p_1\circ p_2\circ\cdots\circ p_m)\circ p_{m+1}\circ\cdots\circ p_{m+n}=p_1\circ p_2\circ\cdots\circ p_{m+n}.$$

但

$$(p_1\circ p_2)\circ p_3=p_1\circ(p_2\circ p_3),\quad p_1\circ p_2=p_2\circ p_1$$

未必成立.

利用这些记号, 上述积分方程可写为

$$e_0=1+\lambda e_0\circ m\circ s, \tag{60.1'}$$

$$e_1=s+\lambda e_1\circ m\circ s. \tag{60.2'}$$

利用逐步逼近法, 形式地得到展开式

$$e_0=1+\lambda m\circ s+\lambda^2 m\circ s\circ m\circ s+\lambda^3 m\circ s\circ m\circ s\circ m\circ s+\cdots,$$

$$e_1=s+\lambda s\circ m\circ s+\lambda^2 s\circ m\circ s\circ m\circ s+\lambda^3 s\circ m\circ s\circ m\circ s\circ m\circ s+\cdots.$$

如能证明这些展开式收敛, 即得上述两个方程的解.

今令

$$\sigma=m\circ s,\quad \mu=s\circ m.$$

上节所述的 σ_2 与 μ_2 就是 $\sigma(r_2),\mu(r_2)$, 因此随着 r_2 为正则边界、流出边界、流入边界与自然边界, $\sigma(r_2),\mu(r_2)$ 就应分别是二者都有限、只前者有限、只后者有限与二者都无限.

由数学归纳法得

$$0\leqslant \underset{(1)}{m\circ} s\circ \underset{(2)}{m\circ} s\circ\cdots\circ \underset{(n)}{m\circ} s(x)\leqslant\frac{\sigma(x)^n}{n!},$$

$$0\leqslant s\circ \underset{(1)}{m\circ} s\circ \underset{(2)}{m\circ} s\circ\cdots\circ \underset{(n)}{m\circ} s(x)\leqslant\frac{s(x)\sigma(x)^n}{n!}.$$

故 e_0,e_1 的级数收敛, 而且有

$$e_0(x)\leqslant \mathrm{e}^{\lambda\sigma(x)}, \tag{60.3}$$

$$e_1(x)\geqslant s(x)\cdot\mathrm{e}^{\lambda\sigma(x)}. \tag{60.4}$$

又有

$$e_0(x) \geqslant \lambda m \circ s(x) = \lambda\sigma(x), \tag{60.5}$$

$$e_1(x) \geqslant \lambda s \circ m \circ s(x). \tag{60.6}$$

现在利用这些不等式来研究 $e_i(x)$ 在 r_2 的极限.

首先注意下列事实:

$$r_2\text{为正则边界} \to m(r_2) < \infty,\ s(r_2) < \infty,$$

$$r_2\text{为流出边界} \to m(r_2) = \infty,\ s(r_2) < \infty,$$

$$r_2\text{为流入边界} \to m(r_2) < \infty,\ s(r_2) = \infty,$$

$$r_2\text{为自然边界} \to m(r_2) \leqslant \infty,\ s(r_2) \leqslant \infty, \text{至少有一个为}\infty,$$

因为

$$\sigma(r_2) = \int_0^{r_2} m(y)\mathrm{d}s(y),$$

所以

$$\sigma(r_2) \geqslant m(r_2')(s(r_2) - s(r_2')),$$

故若 $s(r_2) = \infty$, 则 $\sigma(r_2) = \infty$, 因此当 r_2 为正则边界或流出边界时, $s(r_2) < \infty$. 同样, 当 r_2 为正则边界或流入边界时, $m(r_2) < \infty$.

又由

$$\mu(r_2) = \int_0^{r_2} s(y)\mathrm{d}m(y)$$

有 $\mu(r_2) \leqslant s(r_2) \cdot m(r_2)$. 流出时 $s(r_2) < \infty, \mu(r_2) = \infty$. 故得 $m(r_2) = \infty$. 同样, 流入时 $s(r_2) = \infty$. 若 $m(r_2)$ 与 $s(r_2)$ 均有限, 则 $\mu(r_2), \sigma(r_2)$ 也均有限, 故 r_2 为正则. 自然边界时, m 和 s 中至少有一个为 ∞. 作为 $m(r_2) = \infty, s(r_2) < \infty$, 且 r_2 为自然边界的例子, 可取

$$r_2 = 1,\ s(x) = x,\ m(x) = (1 - x)^{-1}.$$

其次, 作为 $m(r_2) < \infty, s(r_2) = \infty$, 而且 r_2 是自然边界的例子, 可举

$$r_2 = 1,\ s(x) = (1 - x)^{-1},\ m(x) = x.$$

最后, 如令

$$r_2 = \infty,\ s(x) = x,\ m(x) = x,$$

则得到 $m(r_2)=\infty, s(r_2)=\infty, r_2$ 是自然边界的例子.

当 r_2 为正则边界或流出边界时, 由 (60.3) ~ (60.4), $\sigma(r_2)<\infty, s(r_2)<\infty$, 得 $e_0(r_2)<\infty$, $e_1(r_2)<\infty$. 当 r_2 为流入边界或自然边界时, 则 $\sigma(r_2)=\infty$. 故由 (60.5) 得

$$e_0(r_2)=\infty,$$

至于 $e_1(x)$, 由 (60.6) 有

$$\begin{aligned} e_1(x) \geqslant & \lambda s\circ m\circ s(x)=\lambda\int_0^x\int_0^\xi s(\eta)\mathrm{d}m(\eta)\mathrm{d}s(\xi) \\ \geqslant & \lambda s(r_2')\int_{r_2'}^x\int_{r_2'}^\xi \mathrm{d}m(\eta)\mathrm{d}s(\xi)\to\lambda s(r_2')\cdot\sigma(r_2)=\infty. \end{aligned}$$

其次, 为求 $D_s^+e_0(x), D_s^+e_1(x)$, 若形式地微分上述 e_0 与 e_1 的展开式, 得

$$\begin{aligned} D_s^+e_0(x)= & \lambda m+\lambda^2 m\circ s\circ m+\lambda^3 m\circ s\circ m\circ s\circ m+\cdots, \\ D_s^+e_1(x)= & 1+\lambda s\circ m+\lambda^2 s\circ m\circ s\circ m+\cdots. \end{aligned}$$

若能证此式在 I 内为广义一致收敛, 则得证这等式成立. 证法与前面 e_0, e_1 的展开式相同. 而且 $D_s^+e_i(r_2)(i=1,2)$

在r_2为正则边界或流入边界时, 为有限;

在r_2为流出边界或自然边界时, 为无限.

证明也与 $e_i(r_2)$ 的证明相同.

§61 齐次方程 $(\lambda-D_mD_s^+)u=0(\lambda>0)$ 的一般解

令 u_1 和 u_2 为

$$(\lambda-D_mD_s^+)u=0\quad(\lambda>0) \tag{61.1}$$

的两个任意解. 对 u_1, u_2, 称

$$W=W(u_1,u_2)=D_s^+u_1\cdot u_2-D_s^+u_2\cdot u_1$$

为 u_1, u_2 的朗斯基 (Wronskian).

定理 1 W 与 x 无关.

引理 1 令 u, v 为有界变差函数 (function of bounded variation), 则

$$\mathrm{d}(u \cdot v) = v^* \mathrm{d}u + u_* \mathrm{d}v = v_* \mathrm{d}u + u^* \mathrm{d}v,$$

($\mathrm{d}w$ 是对应于 w 的带符号测度 (signed measure), $w^*(x) = w(x+0), w_*(x) = w(x-0)$.)

这引理由

$$\begin{aligned} & u(x_i)v(x_i) - u(x_{i-1})v(x_{i-1}) \\ =& v(x_i)(u(x_i) - u(x_{i-1})) + u(x_{i-1})(v(x_i) - v(x_{i-1})) \\ =& v(x_{i-1})(u(x_i) - u(x_{i-1})) + u(x_i)(v(x_i) - v(x_{i-1})) \end{aligned}$$

而明显. 利用此引理来证明本节定理 1.

$$\begin{aligned} \mathrm{d}W =& u_2 \mathrm{d}(D_s^+ u_1) + D_s^+ u_1 \mathrm{d}u_2 - u_1 \mathrm{d}(D_s^+ u_2) - D_s^+ u_2 \mathrm{d}u_1 \\ =& u_2 \lambda u_1 \mathrm{d}m + D_s^+ u_1 D_s^+ u_2 \mathrm{d}s - u_1 \lambda u_2 \mathrm{d}m - D_s^+ u_2 D_s^+ u_1 \mathrm{d}s = 0. \end{aligned}$$

特别地, 求前节中 e_1, e_0 的朗斯基, 则得

$$W(e_1, e_0) = W(e_1, e_0)(0) = (D_s^+ e_1 \cdot e_0 - D_s^+ e_0 \cdot e_1)(0) = 1.$$

引理 2 $\frac{e_0}{e_1} > \frac{D_s^+ e_0}{D_s^+ e_1} \quad (x > 0)$.

证明 $\frac{e_0}{e_1} - \frac{D_s^+ e_0}{D_s^+ e_1} = \frac{W(e_1, e_0)}{e_1 D_s^+ e_1} = \frac{1}{e_1 D_s^+ e_1} > 0$.

引理 3 $\frac{e_0}{e_1}$ 随 x 增加而减少.

证明 $D_s^+ \left[\frac{e_0}{e_1}\right] = \frac{-W(e_1, e_0)}{e_1^2} = -\frac{1}{e_1^2} < 0$.

引理 4 $\frac{D_s^+ e_0}{D_s^+ e_1}$ 随 x 增加而增加.

证明 $D_m \left(\frac{D_s^+ e_0}{D_s^+ e_1}\right) = \frac{D_m D_s^+ e_0 \cdot D_s^+ e_1 - D_m D_s^+ e_1 \cdot D_s^+ e_0}{(D_s^+ e_1)^2} = \frac{\lambda W(e_1, e_0)}{(D_s^+ e_1)^2} = \frac{\lambda}{(D_s^+ e_1)^2} > 0$.

引理 5 当 $x \uparrow r_2$ 时,

$$\frac{e_0}{e_1} - \frac{D_s^+ e_0}{D_s^+ e_1} = \frac{1}{e_1 D_s^+ e_1} \quad \begin{array}{l} \to 0 \ (\text{当}r_2\text{不是正则边界时}), \\ \to c > 0 \ (\text{当}r_2\text{是正则边界时}). \end{array}$$

有了这些准备后, 试在齐次方程 (61.1) 的解中找出满足二条件:

(A) $u > 0$,

(B) $u \downarrow$ ($x \uparrow$ 时)

的解. 由 (A) 得 $u(0) > 0$, 故求出 $u(0) = 1$ 的解后, 再乘以正的倍数即可. 也就是要求形为

$$u = e_0 - \gamma e_1$$

的解. 由条件 (A),

$$\gamma < \frac{e_0}{e_1} \quad (x > 0)$$

是必要的. 又由 (B) 得 $D_s^+ u < 0$, 故

$$\gamma > \frac{D_s^+ e_0}{D_s^+ e_1} \quad (x > 0)$$

是必要的. 由上述引理,

$$\overline{\gamma} = \lim_{x \uparrow r_2} \frac{e_0}{e_1}, \quad \underline{\gamma} = \lim_{x \uparrow r_2} \frac{D_s^+ e_0}{D_s^+ e_1} \tag{61.2}$$

存在, 而且 γ 满足

$$\overline{\gamma} \geqslant \gamma \geqslant \underline{\gamma}. \tag{61.3}$$

反之, 对如此的 γ, 令

$$u = e_0 - \gamma e_1,$$

则在 $x > 0$ 时, 满足 (A), (B). 至于在 $x < 0$ 时, 只要对 $x < 0$ 考察 e_0, e_1 的展开式就可看出: $e_0, -e$ 都恒为正, 而且是减函数. 综合上述即得下述定理.

定理 2 齐次方程的正而且下降的解, 在 $u(0) = 1$ 的条件下, 由

$$u = e_0 - \gamma e_1$$

给出. 这里 γ 是 $\underline{\gamma}$ 与 $\overline{\gamma}$ 之间的任意正数. 若 γ_2 正则, 则 $\underline{\gamma} < \overline{\gamma}$, 故所求的解有无数个, 都处在 $\overline{u} = e_0 - \underline{\gamma} e_1$ 与 $\underline{u} = e_0 - \overline{\gamma} e_1$ 之间. 在其他场合, $\underline{\gamma} = \overline{\gamma}$, 故 u 是唯一的.

其次, 研究 u 在 r_2 上的值. 若 r_2 为正则边界, 则

$$\begin{aligned}\overline{u}(r_2) =& e_0(r_2) - \frac{D_s^+ e_0(r_2)}{D_s^+ e_1(r_2)} e_1(r_2) = \frac{1}{D_s^+ e_1(r_2)},\\ \underline{u}(r_2) =& e_0(r_2) - \frac{e_0(r_2)}{e_1(r_2)} e_1(r_2) = 0.\end{aligned}$$

对流出边界, 也有 $u(r_2) = 0$.
对流入边界, 则

$$\begin{aligned}u(r_2) &= \lim_{x\to r_2} \left\{ e_0(x) - \frac{D_s^+ e_0}{D_s^+ e_1}(r_2) e_1(x) \right\}\\ &\leqslant \lim_{x\to r_2} \left\{ e_0(x) - \frac{D_s^+ e_0(x)}{D_s^+ e_1(x)} e_1(x) \right\} = \frac{1}{D_s^+ e_1(r_2)}.\end{aligned}$$

又对充分大的 x,

$$u(r_2) + \varepsilon > e_0(x) - \frac{D_s^+ e_0}{D_s^+ e_1}(r_2) e_1(x).$$

其次对充分大的 y,

$$u(r_2) + 2\varepsilon > e_0(x) - \frac{D_s^+ e_0(y)}{D_s^+ e_1(y)} e_1(x).$$

若 $y > x$, 则

$$> e_1(y) - \frac{D_s^+ e_0(y)}{D_s^+ e_1(y)} e_1(y)$$

$$\left(\text{因为} D_s^+ \left\{ e_0(x) - \frac{D_s^+ e_0(y)}{D_s^+ e_1(y)} e_1(x) \right\} = \left\{ \frac{D_s^+ e_0(x)}{D_s^+ e_1(x)} - \frac{D_s^+ e_0(y)}{D_s^+ e_1(y)} \right\} D_s^+ e_1(x) > 0 \right).$$

故

$$u(r_2) + 2\varepsilon > \frac{1}{D_s^+ e_1(y)},$$

因而

$$u(r_2) \geqslant \frac{1}{D_s^+ e_1(r_2)},$$

这样就有

$$u(r_2) = \frac{1}{D_s^+ e_1(r_2)}.$$

对自然边界, 上式虽也成立, 但因 $D_s^+e_1(r_2)=\infty$, 故 $u(r_2)=0$.

用同样的方法也可求得 $D_s^+u(r_2)$. 概括起来就得下述定理.

定理 3　定理 2 的 u 及其 D_s^+u 在 r_2 上的极限值是

$$u(r_2)=\begin{cases} 0 & \begin{array}{l}(\text{当 } r_2 \text{ 为流出边界或自然边界;}\\ \ \text{为正则边界时, 则}\underline{u}=0),\end{array} \\ \dfrac{1}{D_s^+e_1(r_2)} & \begin{array}{l}(\text{当 } r_2 \text{ 为流入边界;}\\ \ \text{为正则边界时, 则}\overline{u}=\dfrac{1}{D_s^+e_1(r_2)}).\end{array}\end{cases}$$

$$D_s^+u(r_2)=\begin{cases} 0 & \begin{array}{l}(\text{当 } r_2 \text{ 为流入边界或自然边界;}\\ \ \text{为正则边界时, 则}D_s^+\bar{u}=0),\end{array} \\ -\dfrac{1}{e_1(r_2)} & \begin{array}{l}(\text{当 } r_2 \text{ 为流出边界;}\\ \ \text{为正则边界时, 则}D_s^+\underline{u}=-\dfrac{1}{e_1(r_2)}).\end{array}\end{cases}$$

因为此 u 与 e_0 是线性独立的, 故任意的解是这二解的线性组合 $\alpha u+\beta e_0$. 因为 $u(r_2), D_s^+u(r_2)$ 都有限, 故若 $\alpha u+\beta e_0$ 为 ∞, 那么这是来自 e_0. 故得下述定理.

定理 4　上述 u 以外的解 v 为

$$|v(r_2)|<\infty(\text{正则边界, 流出边界}), \qquad =\infty(\text{流入边界, 自然边界}),$$
$$|D_s^+v(r_2)|<\infty(\text{正则边界, 流入边界}), \qquad =\infty(\text{流出边界, 自然边界}).$$

§62　非齐次方程 $(\lambda-D_mD_s^+)g=f(\lambda>0)$ 的解

上节中求出的齐次方程的解 u 是正的单调减少的, 而且在 r_2 上, u 和 D_s^+u 都有有限的极限值. 把它记为 $u_2(x)=u_2(x;\lambda)$. 同理, 正的单调增加的解: $u_1(x)=u_1(x;\lambda)$ 也存在, 而且在 r_1 上有同样的极限值.

如上节所述, $W=W(u_1,u_2)$ 是常数, 若考虑 $W(0)$, 就知 $W>0$, 故如以适当的正因子乘 u_1, 就可使 $W=1$.

若令

$$K(x,y) = K(x,y;\lambda) = \begin{cases} u_1(x)u_2(y), & r_1 < x \leqslant y < r_2, \\ u_2(x)u_1(y), & r_1 < y \leqslant x < r_2, \end{cases}$$

则 K 关于变量 x, y 连续且对称. 又 $K \geqslant 0$ 是显然的.

其次, 为求非齐次方程

$$(\lambda - D_m D_s^+)g = f \quad (\lambda > 0) \tag{62.1}$$

的特解 g_0, 设

$$g_0(x) = K \cdot f(x) = \int_{r_1}^{r_2} K(x,y)f(y)\mathrm{d}m(y).$$

此处假定 f 有界、连续, $g_0(x)$ 有意义 (即上述积分是确定的), 而且 $g_0(x)$ 关于 x 的连续性可由下式看出:

$$\begin{aligned} g_0(x) =& u_2(x)\int_{r_1}^{x+0} f(y)u_1(y)\mathrm{d}m(y) + u_1(x)\int_{x+0}^{r_2} f(y)u_2(y)\mathrm{d}m(y) \\ =& \frac{u_2(x)}{\lambda}\int_{r_1}^{x+0} f(y)\mathrm{d}u_1^+(y) + \frac{u_1(x)}{\lambda}\int_{x+0}^{r_2} f(y)\mathrm{d}u_2^+(y). \end{aligned}$$

此处 u^+ 表示 $D_s^+ u$.

既然积分算子 K 的意义已经确定, 它显然是正的, 而且是线性的.

其次, 为了证明 g_0 满足 (62.1), 利用上节的引理 61.1, 就得

$$\begin{aligned} \mathrm{d}g_0 =& \mathrm{d}u_2\int_{r_1}^{x+0} f(y)u_1(y)\mathrm{d}m(y) + u_2(x)f(x)u_1(x)\mathrm{d}m \\ & + \mathrm{d}u_1\int_{x+0}^{r_2} f(y)u_2(y)\mathrm{d}m(y) - u_1(x)f(x)u_2(x)\mathrm{d}m \\ =& \mathrm{d}u_2\int_{r_1}^{x+0} f(y)u_1(y)\mathrm{d}m(y) + \mathrm{d}u_1\int_{x+0}^{r_2} f(y)u_2(y)\mathrm{d}m(y), \\ D_s^+ g_0 =& D_s^+ u_2\int_{r_1}^{x+0} f(y)u_1(y)\mathrm{d}m(y) + D_s^+ u_1\int_{x+0}^{r_2} f(y)u_2(y)\mathrm{d}m(y). \end{aligned}$$

再一次进行同样的计算, 并利用 $\lambda u_i = D_m D_s^+ u_i$ 就得

$$D_m D_s^+ g_0 = \lambda g_0 - f.$$

当 r_1, r_2 中之一或二者为正则边界时, 对应的 u_1, u_2 有无限多个, 故 K 的选法有无数个, 因而 g_0 也有无数个, 但其中任何一个都有上述性质.

研究 g_0 在两端上的极限值, 则得下面的定理, 应写为 $g_0(r_2-)$ 的记为 $g_0(r_2)$.

定理 1

(i) 若 r_2 为正则边界, 则 $g_0(r_2) = u_2(r_2)\int_{r_1}^{r_2} fu_1\mathrm{d}m$,

$$g_0^+(r_2) = u_2^+(r_2)\int_{r_1}^{r_2} fu_1\mathrm{d}m.$$

(ii) 若 r_2 为流出边界, 则 $g_0(r_2) = 0$.

(iii) 若 r_2 为流入边界, 则 $g_0(r_2) = u_2(r_2)\int_{r_1}^{r_2} fu_1\mathrm{d}m, g_0^+(r_2) = 0$.

(iv) 若 r_2 为自然边界, 则 $f(r_2)$ 存在时, $g_0(r_2)$ 也存在, 且

$$g_0(r_2) = f(r_2)/\lambda.$$

证明 (i) 是明显的, 下面证 (ii).

由于

$$|g_0(x)| \leqslant (K\cdot 1)(x)\sup_x |f(x)|,$$

故只要证明 $(K\cdot 1)(x) \to 0 (x \to r_2)$ 即可.

$$\begin{aligned}
\lambda(K\cdot 1)(x) =& u_2(x)\int_{r_2}^{x+0} \mathrm{d}u_1^+ + u_1(x)\int_{x+0}^{r_2} \mathrm{d}u_2^+ \\
=& u_2(x)u_1^+(x) - u_2(x)u_1^+(r_1) + u_1(x)[u_2^+(r_2) - u_2^+(x)] \\
=& u_2(x)u_1^+(x) + o(1) = u_1^+(x)\int_x^{r_2} (-u_2^+)\mathrm{d}s + o(1) \\
\leqslant& -u_2^+(x)\int_x^{r_2} u_1^+(y)\mathrm{d}s + o(1) \quad (\text{因为} -u_2^+ \downarrow, u_1^+ \uparrow) \\
=& -u_2^+(x)(u_1(r_2) - u_1(x)) + o(1) \to 0.
\end{aligned}$$

(iii) 的证明. 令 $M = \sup\limits_{r_1\leqslant y\leqslant r_2} |f(y)|$, 由 $u_2^+(r_2) = 0$ 得

$$\left|u_1(x)\int_{x+0}^{r_2} f(y)\mathrm{d}u_2^+(y)\right| \leqslant -Mu_1(x)u_2^+(x) = M(1 - u_1^+(x)u_2(x)).$$

另一方面, 注意到存在某 $\widetilde{\gamma} > 0$, 使 u_1 有如下表示,

$$u_1(x) = \frac{1}{\gamma + \widetilde{\gamma}}(e_0(x) + \widetilde{\gamma}e_1(x)).$$

因此,

$$u_2(r_2)u_1^+(r_2) = \frac{1}{\gamma + \widetilde{\gamma}}\left\{\frac{e_0^+(r_2)}{e_1^+(r_2)} + \widetilde{\gamma}\right\} = 1,$$

$$\lim_{x \to r_2} u_1(x)\int_{x+0}^{r_2} f(y)\mathrm{d}u_2^+(y) = 0.$$

(iii) 的后半部分显然成立.

(iv) 的证明. 首先注意 $u_2(r_2) = 0, u_2^+(r_2) = 0$. 对任意 $\varepsilon > 0$, 存在 ξ, 使得对任意 $x > \xi$, 有 $|f(x) - f(r_2)| < \varepsilon$. 因此, 在 $x > \xi$ 时, 有

$$\begin{aligned} g_0(x) =& u_2(x)\int_{r_1}^{\xi} f(y)u_1(y)\mathrm{d}m(y) \\ &+ \frac{u_2(x)}{\lambda}\int_{\xi}^{x+0} f(y)\mathrm{d}u_1^+(y) + \frac{u_1(x)}{\lambda}\int_{x+0}^{r_2} f(y)\mathrm{d}u_2^+(y) \\ \leqslant& u_2(x)\int_{r_1}^{\xi} f(y)u_1(y)\mathrm{d}m(y) \\ &+ \frac{f(r_2)+\varepsilon}{\lambda}(u_2(x)u_1^+(x) - u_2(x)u_1^+(\xi) - u_1(x)u_2^+(x)) \\ =& u_2(x)\int_{r_1}^{\xi} f(y)u_1(y)\mathrm{d}m(y) + \frac{f(r_2)+\varepsilon}{\lambda}(1 - u_2(x)u_1^+(\xi)). \end{aligned}$$

令 $x \to r_2$ 可得 $g_0(r_2) \leqslant \frac{f(r_2)+\varepsilon}{\lambda}$. 同理有 $g_0(r_2) \geqslant \frac{f(r_2)-\varepsilon}{\lambda}$.

例 设 $I = (-\infty, +\infty), D_m D_s^+ = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}$. 则 $-\infty, +\infty$ 都是自然边界, 故 u_1, u_2 除常数因子外唯一确定. 解

$$\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}u_i = \lambda u_i, \quad u_i > 0, \quad u_1 \uparrow,\ u_2 \downarrow,$$

而得

$$u_1 = \mathrm{e}^{\sqrt{2\lambda}x}, \quad u_2 = \mathrm{e}^{-\sqrt{2\lambda}x}.$$

因为 $W(u_1, u_2) = 2\sqrt{2\lambda}$, 故若以 $u_1/(2\sqrt{2\lambda})$ 代替 u_1, 则得 $W = 1$, 故有

$$K = \frac{1}{2\sqrt{2\lambda}}\mathrm{e}^{-\sqrt{2\lambda}|x-y|}.$$

因为 $s = x, m = 2x$, 故

$$Kf = \frac{1}{\sqrt{2\lambda}}\int_{-\infty}^{\infty} \mathrm{e}^{-\sqrt{2\lambda}|x-y|} f(y)\mathrm{d}y.$$

§63　$x^{(a)}(t)$ 诸量在正则区间上的分布

至今为止, 只研究了 Feller 算子 $D_mD_s^+$ 的解析性质, 现在再来研究原来的问题. 令 $x^{(a)}(t)$ 为在紧致的 (compact) 距离空间 R 上变动的 Markoff 过程, 令 R 的开集 I 为正则区间. 也就是说, 令 I 与实数区间 (r_1, r_2) 同胚, I 的各点不仅是扩散点, 而且还是正则点. 前已说过, 在 I 上存在典范尺度 s 和典范测度 $\mathrm{d}m$, 使

$$A_I = (D_mD_s^+)_I. \tag{63.1}$$

像以前一样, 把 I 的点与 (r_1, r_2) 的点同样看待. 以 $P_I(t, \xi, E)$ 表示从 I 的点 ξ 出发未跑出 I 而经 t 时后到达 E 的概率. 若令 $\tau_I^{(\xi)}$ 为 $x^{(\xi)}(t)$ 跑出 I 的最初时间, 则

$$P_I(t, \xi, E) = p\{x^{(\xi)}(t) \in E,\ \tau_I^{(\xi)} > t\}. \tag{63.2}$$

当 $\tau_I^{(\xi)} < \infty$ 时, $x^{(\xi)}(\tau_I^{(\xi)})$**不属于**$I$ 本身, 而属于 I 在 R 中的边界点上, 但对任意 $\delta > 0, x^{(\xi)}(\tau_I^{(\xi)} - \delta)$ 则属于 I, 而且由于 $x^{\xi}(t)$ 的样本过程连续, 故 $x^{\xi}(\tau_I^{\xi} - 0)$ 确定而且等于 r_1 或 r_2. 令 $\tau_1^{(\xi)} = \tau_1^{(\xi)}(I)$, 在 $\tau_I^{(\xi)} < \infty$ 并且 $x^{(\xi)}(\tau_I^{(\xi)} - 0) = r_1$ 时等于 $\tau_I^{(\xi)}$, 在其他情形则等于 ∞. 同样利用 r_2 来定义 $\tau_2^{(\xi)} = \tau_2^{(\xi)}(I)$. $\tau_i^{(\xi)}$ 就是 $x^{(\xi)}(t)$ 从 r_i 这边跑出 I 的时间. $\tau_i^{(\xi)}$ 的分布记为 $\varphi_i(\mathrm{d}t, \xi)$.

不求 $P_I(t, x, E), \varphi_i(\mathrm{d}t, \xi)$ 而计算它们的 Laplace 变换

$$R_I(\lambda, x, E) = \int_0^\infty \mathrm{e}^{-\lambda t} P_I(t, x, E)\mathrm{d}t, \tag{63.3}$$

$$\hat{\varphi}_{i\lambda}(\xi) = \int_0^\infty \mathrm{e}^{-\lambda t} \varphi_i(\mathrm{d}t, \xi). \tag{63.4}$$

本节的目的就是说明下面的一些定理.

定理 1　利用上节的 $u_{i\lambda}(\xi)$ $(i = 1, 2)$ 及 K_λ 就得

$$R_I(\lambda, x, E) = \int_E K_\lambda(x, y) m(\mathrm{d}y), \tag{63.5}$$

$$\hat{\varphi}_{1\lambda}(\xi) = \frac{u_{2\lambda}(\xi)}{u_{2\lambda}(r_1)}, \quad \hat{\varphi}_{2\lambda}(\xi) = \frac{u_{1\lambda}(\xi)}{u_{1\lambda}(r_2)}. \tag{63.6}$$

若 r_i 为正则边界时, $u_{i\lambda}(\xi)$ 应取为 $\underline{u}_{i\lambda}(\xi)$, 即 $u_{i\lambda}(r_i)=0$ 的解, 而 $K_\lambda(x,y)$ 则由此等解构成.

在 (63.6) 中, 譬如, $u_{2\lambda}(r_1)=\infty$, 则 $\hat{\varphi}_{1\lambda}(\xi)=0$, 即 $P(\tau_1^{(\xi)}=\infty)=1$, 这表示不能到达 r_1. 若 $u_{2\lambda}(r_1)<\infty$, 则 $\hat{\varphi}_{1\lambda}(\xi)>0$, 且 $P(\tau_1^{(\xi)}<\infty)>0$, 而表示有到达 r_1 的可能性. 因此得下述定理

定理 2 在有限时间内可能到达正则边界与流出边界, 但在有限时间内不能到达流入边界和自然边界.

今证定理 1. 先在 I 内取区间 $J=(j_1,j_2)(\overline{J}\subset I)$, 而对 J 证明定理. 对 J 采用 $R_J(\lambda,x,E), K_\lambda(x,y;J), \varphi_i(\mathrm{d}t,\xi,J), \hat{\varphi}_{i\lambda}(\xi;J), u_{i\lambda}(\xi;J)$, 等记号. 如前一样, 利用强 Markoff 性, 对 $f\in \boldsymbol{C}(R)$ 得

$$R_\lambda f(x)=R_{J\lambda}f(x)+\sum_{i=1}^{2}\hat{\varphi}_{i\lambda}(x,J)R_\lambda f(j_i). \tag{63.7}$$

因为 $\hat{\varphi}_{1\lambda}(x,J)$ 是 $(\lambda-A_J)u=0, u(j_1)=1, u(j_2)=0$ 的解, 故由 $A_J=(D_mD_s^+)_J$ 得

$$\hat{\varphi}_{1\lambda}(x,J)=\frac{u_{2\lambda}(\xi;J)}{u_{2\lambda}(j_1;J)}. \tag{63.8}$$

同理

$$\hat{\varphi}_{2\lambda}(x,J)=\frac{u_{1\lambda}(\xi;J)}{u_{1\lambda}(j_2;J)}. \tag{63.9}$$

因而

$$R_\lambda f(x)=R_{J\lambda}f(x)+u_{1\lambda}(\xi;J)\frac{R_\lambda f(j_2)}{u_{1\lambda}(j_2;J)}+u_{2\lambda}(\xi;J)\frac{R_\lambda f(j_1)}{u_{2\lambda}(j_1;J)}. \tag{63.10}$$

另一方面, 又因 $R_\lambda f$ 满足 $(\lambda-A)R_\lambda f=f$, 显然在 J 内有 $(\lambda-A_J)R_\lambda f=f$. 由 $A_J=(D_mD_s^+)_J$ 且用前节的结果, 就有

$$R_\lambda f(x)=K_{\lambda J}f(x)+c_1u_{1\lambda}(x;J)+c_2u_{2\lambda}(x;J),\ x\in J,$$

因为 J 的两端是正则边界, 故

$$K_{\lambda J}(j_i)=0,\ u_{i\lambda}(j_i;J)=0,\ i=1,2,$$

因而

$$c_1=\frac{R_\lambda f(j_2)}{u_{1\lambda}(j_2;J)},\quad c_2=\frac{R_\lambda f(j_1)}{u_{2\lambda}(j_1;J)}.$$

故

$$R_\lambda f(x)=K_{\lambda J}f(x)+\frac{R_\lambda f(j_2)}{u_{1\lambda}(j_2;J)}u_{1\lambda}(x;J)+\frac{R_\lambda f(j_1)}{u_{2\lambda}(j_1;J)}u_{2\lambda}(x;J). \quad (63.11)$$

比较 (63.10) 与 (63.11) 得

$$R_{J\lambda}f(x)=K_{\lambda J}f(x),\quad x\in J, \quad (63.12)$$

因任意在 $\overline{J}$ 上连续的函数都可推广为 R 全体上的连续函数, 故 (63.12) 对所有的 $f\in C(J)$ 成立, 因此得

$$R_J(\lambda;x,dy)=K_J(\lambda,x,y)\mathrm{d}m(y)=\begin{cases}u_{1\lambda}(x;J)u_{2\lambda}(y;J)\mathrm{d}m(y), & j_1\leqslant x\leqslant y\leqslant j_2,\\ u_{2\lambda}(x;J)u_{1\lambda}(y;J)\mathrm{d}m(y), & j_1\leqslant y\leqslant x\leqslant j_2.\end{cases}$$

若令 $J\uparrow I$, 则 $\tau_J^{(\xi)}\uparrow\tau_I^{(\xi)}$, 而且由

$$R_{J\lambda}f(\xi)=E\left\{\int_0^{\tau_J^{(\xi)}}\mathrm{e}^{-\lambda t}f(x(t))\mathrm{d}t\right\},$$

故 $R_{J\lambda}f(\xi)\to R_{I\lambda}f(\xi), u_{1\lambda}(x;J)$ 和 $u_{2\lambda}(x;J)$ 的选择虽只有常数因子的自由度, 但当 $J\uparrow I$ 时, 可以分别使之接近于 $u_{1\lambda}(x)$ 和 $u_{2\lambda}(x)$, 因此, $K_{\lambda J}f(\xi)\to K_\lambda f(\xi)$, 并得 $R_{I\lambda}=K_{\lambda I}$, 故得 (63.5), 至于 $\hat{\varphi}_{i\lambda}(\xi)$, 当 $J\uparrow I$ 时, 先令 $j_2\uparrow r_2$, 其次令 $j_1\downarrow r_1$, 则易由 (63.8 ~ 63.9) 推出 (63.6).

§64 在正则区间的边界上的行动

采用上节的记号. 因为 I 与 (r_1,r_2) 同胚, 故对于无论多么接近于 r_1 的实数, 都有对应的 I 点, 但对 r_1 本身就没有对应的点. 若 $\tau_I^{(\xi)}$ 有限, 则 $x(\tau_I^{(\xi)})$ 是 I 在 R 的边界点. 若令 b 为 I 的边界上的一点, 则存在逼近于 b 的 I 的点列 b_n. 对应于 b_n 的实数 (b_n 的坐标) 序列以 r_1,r_2 中之一或二者作为极限点. 因此可以假定以二者之一作为极限点. 若 b_n 以 r_i 作为极限点, 就是" b 对应于 r_i". 一点 b 可能对应于 r_1,r_2 双方, 又不同的 b 与 b' 也可能对应于同一 r_i. 后者仅当 r_i 是自然边界时才有可能 (见本节定理 1). 今若令 b 与 b' 对应于 r_1, 则得 $b_n\to b, b_n'\to b', \overline{b}_n(b_n$ 的坐标$)\to r_1, \overline{b'}_n\to r_1$. 如前节所述, 在 I 内有

$$R_\lambda f = K_\lambda f + c_1 u_1 + c_2 u_2.$$

当 r_1 为流入边界时, 则 $u_2(r_1+) = \infty$, 故为使上式成立就必须有 $c_2 = 0$. 因为 $K_\lambda f(r_1+), u_1(r_1+)$ 都存在, 所以

$$R_\lambda f(b) = \lim_n R_\lambda f(b_n) = \lim_n \{K_\lambda f(\bar{b}_n) + c_1 u_1(\bar{b}_n)\} = K_\lambda f(r_1+) + c_1 u_1(r_1+).$$

同理, $R_\lambda f(b') = K_\lambda f(r_1+) + c_1 u_1(r_1+) = R_\lambda f(b)$. 由 $\lambda R_\lambda f \to f$, 得 $f(b') = f(b)$. 因为 f 是 $\boldsymbol{C}(R)$ 中任意的元, 故得 $b = b'$. 又 r_1 为流出边界或正则边界时, $u_2(r_1+)$ 也有限地确定, 故如上一样得 $b = b'$. 从而有下述定理.

定理 1 若 r_i 不是自然边界, 则对应于 r_i 的 I 的边界点只有一个.

(i) 研究 $x^{(a)}(t)$ 在对应于自然边界的边界点上的行动. 令 r_1 为自然边界, 那么与此对应的边界点一般地有许多个, 并且构成 R 的闭子集 F. 又从 F 出发的 $x^{(a)}(t)$ 恒被关闭在 F 内. 为此, 考虑在 F 的邻域 U 上为 0 的 R 上的连续函数 f, 则 f 在 I 上的 r_1 的邻域上为 0. 故 $f(r_1-)$ 存在而且当然为 0, 于是 $R_\lambda f(r_1-) = f(r_1-)/\lambda = 0$. 因而得 $R_\lambda f(a) = 0, a \in F$. 今若取 f 在 F 上为 0, 在 U 外为 1, 在 $U - F$ 上则位于 0, 1 之间, 则

$$R_\lambda f(a) \geqslant \int_0^\infty \mathrm{e}^{-\lambda t} P(t, a, U^c) \mathrm{d}t.$$

因为左边为 0, 故 $P(t, a, U^c) = 0$, 亦即 $P(t, a, U) = 1$. 令 $U \downarrow F$, 就得

$$P(t, a, F) = 1.$$

因为 $x^{(a)}(t)$ 连续, 故

$$P(x^{(a)}(t) \in F, \quad 0 \leqslant t < \infty) = 1, \quad a \in F.$$

特别地, 若 F 是一点, 则此点就是套点. 一般从 I 的点出发, 在有限时间内不可能到达 F.

(ii) 流出边界上的状态. 令 r_1 为流出边界. 因为对应于 r_1 的边界仅有一个, 故也用 r_1 表示. 若取 r_1 的邻域 U 而且考虑其边界, 则得 I 内的一点 r_1' 及不在 I 中的闭集 C. 从 U 中的一点 a 出发, 记在到达 C 以前先到达 r_1' 的时间为 $\tau^{(a)}$, 若这种情况不发生, 则令 $\tau^{(a)} = \infty$. 令

$$U(a) = E\{\mathrm{e}^{-\lambda \tau^{(a)}}\},$$

若 a 处于 (r_1, r_1') 内, 则

$$(\lambda - D_m D_s^+)U = 0,$$

$$U(r_1') = 1.$$

若 $r_1 \leqslant x < y < r_1'$, 则 $U(x) \leqslant U(y)$.

故得

$$U(r_1) \leqslant U(x) = \frac{u_{1\lambda}(x)}{u_{1\lambda}(r_1')} \quad (r_1 \leqslant x < r_1').$$

若 r_1 为流出边界, 则 $u_{1\lambda}(x) \to 0$ $(x \to r_1)$. 故得 $U(r_1) = 0$. 这表示若从 r_1 出发, 则不能自 r_1 这边进入 I 内. 如前所述, 从 I 这边到达 r_1 是可能的. 这就是称 r_1 为流出边界的原因.

(iii) 考虑流入边界. 早已看到从 I 这边不能到达流入边界. 对应于流入边界 r_1 的 I 在 R 上的边界点只有一个, 也以 r_1 表示. 若 r_1 在 R 内为 $x^{(a)}(t)$ 的扩散点, 则自 r_1 出发立即进入 I, 然后从 I 这边就不能到达 r_1. 这就是称 r_1 为流入边界的原因.

首先证明 r_1 不是套点. 取在 R 上的连续函数 f, 令 $f \geqslant 0, f(r_1) = 0, f(r_2') = 1$($r_2'$ 是 I 内的一点). 当 $r_1 < a < r_2'$ 时,

$$R_\lambda f(a) \geqslant \hat{\varphi}_{2\lambda}(a) R_\lambda f(r_2').$$

这里 $\hat{\varphi}_{2\lambda}(a)$ 是对区间 (r_1, r_2') 而取的. 故它等于 $u_{1\lambda}(r_2')/u_{1\lambda}(a)$, 而且当 $a \to r_1$ 时逼近于 $u_{1\lambda}(r_2')/u_{1\lambda}(r_1)$. 故得

$$R_\lambda f(r_1) \geqslant \frac{u_{1\lambda}(r_2')}{u_{1\lambda}(r_1)} \cdot R_\lambda f(r_2').$$

若 r_1 是套点, 则左边等于 $f(r_1)/\lambda = 0$, 但由 $\lambda R_\lambda f(r_2') \to f(r_2') = 1$ 可以推知, 对足够大的 λ 右边是正的. 于是得出矛盾, 故 r_1 不是套点. 进一步还可证明, 从 r_1 出发一定立即进入 I 内, 但因证明复杂, 故从略. 由此可知 r_1 是一般右通过点.

今研究在 r_1 上的生成算子. 取 r_1 的充分小的邻域 U, 使 $p_U(\xi) = E(\tau_U^{(\xi)}) < \infty$. 若以 $\tau(\xi)$ 表示从 r_1 到达 ξ 的最初时间, 则因 r_1 是通过点, 故要跑出 U, 就必须通过 ξ, 因而 $\tau(\xi) < \tau_U^{(r_1)}$. 若令 $p(\xi) = E(\tau(\xi))$, 则

$$p_U(r_1) = p(\xi) + p_U(\xi).$$

但由引理 55.2,

$$A_J p_U(\xi) = -1 \quad (J = I \cap U).$$

故 $A_J p(\xi) = 1$. 亦即 $D_m D_s^+ p(\xi) = 1$. 由此得

$$p(\xi) = \int_{\xi_0}^{\xi} m(x) \mathrm{d}s(x) + as(\xi) + b.$$

此处 ξ_0 是比 ξ 小的任意常数. 若取 $\xi_0 = r_1$, 则上面的积分发散, 因而不方便, a, b 与 ξ_0 有关, 只有在 $\xi > \xi_0$ 时上式才妥当. 将上式变形为

$$p(\xi) = \int_{r_1}^{\xi} (s(\xi) - s(x)) \mathrm{d}m(x) + cs(\xi) + d,$$

则它不含 ξ_0, 而且恒成立. 因为 $p(r_1) = 0$, 故不得不有 $c = d = 0$(注意 $s(r_1) = -\infty$). 于是得

$$p(\xi) = \int_{r_1}^{\xi} (s(\xi) - s(x)) \mathrm{d}m(x).$$

因而由 §56 的定理 1,

$$A_{r_1} f = \lim_{\xi \downarrow r_1} \frac{f(\xi) - f(r_1)}{\int_{r_1}^{\xi} (s(\xi) - s(x)) \mathrm{d}m(x)}.$$

(iv) 在正则边界 r_1 上的行动有好几种不同的可能性, 它相应于在正则边界上选择 $u_{1\lambda}(x)$ 的多样性. 这里考虑这样的情况: r_1在 R 内也是端点, 而且 r_1 的邻域只是 r_1 或只是 I 的点. 这时 r_1 或者是套点, 或者是一般右通过点. 若是套点, 则 $R_\lambda f(r_1) = f(r_1)/\lambda$ 故

$$R_\lambda f(x) = K_\lambda f(x) + \frac{u_{2\lambda}(x)}{u_{2\lambda}(r_1)} \cdot \frac{f(r_1)}{\lambda} + \frac{u_{1\lambda}(x)}{u_{1\lambda}(r_2)} \cdot R_\lambda f(r_2).$$

若是一般右通过点, 则与 (iii) 同样求 $p(\xi)$ 而得

$$p(\xi) = \int_{r_1}^{\xi} (m(x) + c) \mathrm{d}s(x).$$

因为 $m(x)$ 的决定有一附加常数的选择自由, 故可令 $c = 0$ 以使

$$p(\xi) = \int_{r_1}^{\xi} m(x) \mathrm{d}s(x).$$

因为 $p(\xi)\uparrow$, 故 $m(x)>0$. 也就是

$$m(r_1)\geqslant 0,$$

而且

$$A_{r_1}f=\lim_{\xi\downarrow r_1}\frac{f(\xi)-f(r_1)}{\int_{r_1}^{\xi}m(x)\mathrm{d}s(x)}.$$

若 f 不但属于 $\mathfrak{D}(A_{r_1})$, 而且属于 $\mathfrak{D}(A_U)$, 这里 U 是 r_1 的邻域, 则 $\xi>r_1$, 而且 $f\in\mathfrak{D}((D_mD_s^+)_J)$, $J=U\cap I$, 故得

$$A_Uf(r_1)=\lim_{\xi\downarrow r_1}\frac{D_s^+f(\xi)}{m(\xi)}.$$

若 $m(r_1)=0$, 则

$$D_s^+f(r_1)=0.$$

这叫做**反射壁**(reflecting barrier) 的边界条件.

若 $m(r_1)>0$, 则由 $A_Uf(x)$ 的连续性得

$$\frac{D_s^+f(r_1)}{m(r_1)}=D_mD_sf(r_1).$$

这叫做**普遍化了的反射壁的边界条件**, 是由 Feller 最先引进的.

若考察上述边界条件的概率意义, 则会发现许多有趣的问题.

后　　记

现在列举本书的参考读物如下.

A. Kolmogoroff: Grundbegriffe der Wahrscheinlichkeitsrechnung, Erg. der Math., Berlin, 1933①.

W. Feller: An introduction to probability theory and its application, 1950②.

J. L. Doob: Stochastic processes, 1952.

P. Lévy [1]: Théorie de l'addition des variables aléatoires, Paris, 1937.

[2]: Processus stochastiques et mouvement brownien, Paris, 1948.

河田敬义: 确率论, 共立出版社, 1948.

国泽清典: 近代确率论, 岩波全书, 岩波书店, 1951.

伊藤　清: 确率论, 现代数学, 岩波书店, 1952.

丸山仪四郎: 确率论, 现代数学讲座, 共立出版社, 1957.

对本书各章附加说明如下.

第 1 章　这一章叙述了概率论的基本概念. 要把概率论构成为一门数学, 最适合的方法是 Kolmogoroff 的测度方法. 本书就是根据这个观点叙述的. 但是要善于在实际问题中加以应用, 那就必须充分理解其直观背景. 从这个角度来看, 上面所列 Feller 的书值得推荐.

对于 Kolmogoroff 的测度方法的基本概念, Kolmogoroff 的书和上面所列河田敬义的书都可参考. 由于这两本书都没有详细地讨论随机过程的样本过程, 为此有必要导入 Doob 的可分性. Doob 原先发表的论文, 观点稍为不彻底, 而且晦涩难读, 但上面所列 Doob 的书已经过整理, 有详尽的阐述.

第 2 章　本章着重讨论时间参数为连续的情形 (可加过程). 对于参

① 中译本《概率论基本概念》, 丁寿田译, 商务印书馆. —— 校者注

② 《概率论及其应用》, 人民邮电出版社出版了本书的中文版, 分卷1和卷2. —— 编者注

数为离散的情形 (可加序列), 所讨论的仅限于研究可加过程所需要的范围之内, 例如大数定律、中心极限定理、重对数定理等都省略, 关于这方面, 可参考上述国泽清典的书. 另外, 钟开莱 (K. L. Chung) 由俄文版译成英文版的下列书中也有很好的阐述.

Gnedenko-Kolmogoroff: Limit distribution for sums of independent random variables, Moscow-Leningrad, 1949①

对于可加过程, 上述 P. Lévy 的 [1] 和 [2] 内容最为丰富. 但是读起来相当困难, 其中的基本部分, 在上述 Doob 及伊藤清的书内, 都根据 Kolmogoroff 的观点进行了整理, 比较容易理解.

第 3 章 本章叙述了平稳过程的基本内容, 不过略去了 Wiener-Kolmogoroff 内插, 外推, 以及 Grenander 的参数估计等知识点. 这将在本丛书的河田龙夫著《确率过程的应用》中阐述, 参考文献如下.

上述 Doob 著作的第 12 章.

N. Wiener: Extrapolation, interpolation and smoothing of stationary time series, 1949.

A. Kolmogoroff: Interpolation und Extrapolation in stationären Zufälligen Folgen, Bull. Acad. Sci. U.R.S.S. Ser. Math., 5, 1941.

U. Grenander: Stochastic processes and statistical inference, Arkiv. för Mat., 1, 1950.

另外, 对于本章提到而不加以证明的调和分析, 可参看

N. Wiener: Generalized harmonic analysis, Acta. Math., 55, 1930.

第 4 章及第 5 章 这两章的主题限于对时间为齐次的 Markoff 过程, 并简称为 Markoff 过程, 理由是仅有这种情形才有完整的理论, 但这里也并未罗列所有的知识点. 第 4 章仅讲一些基本的结果, 第 5 章关于扩散, 添上 Dynkin 方法, 介绍了 Feller 的最新理论. 至于 Markoff 过程的遍历 (ergodic) 性, 由于已经有了很多的研究, 受到篇幅上的限制只好割爱.

记载古典结果的文献有

M.Fréchet: Recherches théoriques modernes sur le calcul des probabilités, second liver, methode des functions arbitraires, théorie des événement en chain dans le cas d'un nombre fini d'états possibles, Paris, 1938.

---

① 中译本《相互独立随机变量和的极限分布》, 王寿仁译, 科学出版社. —— 校者注

如书名所示, 该书仅讲述有限个状态的 Markoff 过程. 关于可数个状态的 Markoff 过程, 创始期的著作是

P. Lévy: Systèms markoviens et stationaires. Cas dénombrable Ann. Sci. École Norm. Sup., 68, 1951.

最近钟开莱发表了很多使这方面问题进一步精密化的研究成果. 这方面的日文参考读物以前面所说的丸山仪四郎的书为最好.

关于时间不是齐次的 Markoff 过程, 也有了不少的研究, 可参看上述 Feller, 伊藤清和 Doob 等人的书. 作为这方面研究的开端, 下列文献是值得推荐的:

A. Kolmogoroff: Analytische Methoden in der Wahrscheinlichkeitsrechnung, Math. Ann., 104.

校 后 记

郑绍濂

随机过程论是概率论中的一个重要分支. 它是在近三十年来, 由于物理学、无线电技术、力学等方面的需要而迅速发展起来的一门学科.

虽然早在 1900 年, 法国数学家巴歇列 (Bachelir) 已对随机过程的理论进行过研究 (Ann.de I'Écode norm 17, 1900, p.21), 但是随机过程的严密理论直到 1931 年出现了前苏联数学家科尔莫戈罗夫 (A. H. Колмогоров) 的著作[28] 之后, 才开始建立. 这篇文章, 开辟了随机过程论中的一个重要方向——马尔可夫过程的研究. 1934 年, 前苏联的另一数学家辛钦 (A. Я. Хинчин) 在他的著作[21] 中第一次提出随机过程论另一重要领域 —— 平稳随机过程的严密理论, 并阐述了它与具有不变测度位相空间的动力学系统的联系. 其后, 有不少数学家在随机过程论的各个领域中进行了大量的工作. 现今, 随机过程的理论不仅可算是近代发展最快的数学分支之一, 而且它的应用范围几乎已涉及一切自然科学与工程技术领域. 特别地, 它已成为研究现代物理、自动控制、无线电技术等先进科学技术的有力工具.

日本数学家伊藤清所著的《随机过程》是一本具有较高理论水平的著作. 作者以简练的方法介绍了随机过程的基本理论, 对于数学基础较好的读者来说, 通过阅读这本书, 可以较快地掌握构成随机过程论的几个主要方面的思想与方法, 从而打下较好的理论基础.

本书共分四个部分. 第一部分 (即本书的第 1 章) 介绍了科尔莫戈罗夫关于随机过程的基本理论; 第二部分 (即本书的第 2 章) 介绍了可加过程的一些基本结果, 着重讨论了可加过程样本函数的性质, 并导出了无穷可分分布律的典型表示式; 在第三部分中 (即本书的第 3 章), 作者以泛函分析为工具, 介绍了平稳随机过程的谱分解理论与遍历理论, 同时, 也扼要介绍了由作者本人首先研究的平稳广义过程; 最后一部分 (即本书的第 4, 5 两章), 也是本书写得最精彩的部分, 作者对研究马尔可夫过程方面的

最新方法 (即半群的方法) 作了透彻的介绍. 由于作者的精心编排, 虽然这一部分所占的篇幅不多, 但从本质上对问题作了阐明.

这本书虽然具有较高的学术水平, 但作为一本随机过程基本理论方面的著作, 还是有一些缺陷. 首先, 对于一些在应用上与理论上都很重要的内容, 如关于平稳过程的预测理论、马尔可夫过程的遍历性理论等, 作者有意识地将它们舍弃了. 其次, 在全书各个重要理论的陈述上, 作者采用的是纯粹的逻辑推理方法, 而对这些理论与实际的联系注意得不够, 如随机游动与质点的勃朗运动的联系、平稳广义过程与勃朗运动中质点的速度、电子学中的白噪声等的联系等, 这些联系正是需要建立平稳广义过程理论的主要的物理依据. 同时, 由于缺少直观的说明, 有些概念 (如第 3, 4 两章的开头部分) 对初学的人来说难于接受. 最后, 所举出的参考文献相对于书中的丰富内容来说, 显得太少了, 这会给读者带来许多的不便.

总的来说, 这本书的翻译出版, 对于我国广大的读者是有所裨益的.

为了使读者便于查阅和深究本书的有关内容, 在校后记中作如下补充是有益的.

阅读本书所需的有关概率论知识, 可在复旦大学数学系主编的《概率论与数理统计》(上海科技出版社出版) 一书中找到.

对于本书要用到的有关测度论的知识, 读者可参看 Paul R. Halmos 著的《测度论》(王建华译, 科学出版社出版). 本书第 3、4、5 章需用到的关于泛函分析方面的知识, 可参看关肇直编著的《泛函分析讲义》(高等教育出版社出版). 第 3 章引用的广义函数论的知识, 除 §31 中关于推广的 Bochner 定理而外, 都可在冯康的著作[6] 中找到.

第 2 章 §17 中对依概率连续的可加过程的样本函数性质的讨论, 可参看 J. R. Kinney 的著作[27]. 对 §16, 还可参看Ю. В. Прохоров[30] 与A. В. Скороход[30] 等人有关对应于可加过程的概率测度的极限理论.

第 3 章 从泛函分析的观点对平稳过程作系统介绍的著作, 当参数为离散时, 可参看A. Н. Кодмогорв的著作[20]. 当参数为连续时, 可参看 K. Karhunen 的著作[24][25]. 还可参看A. М. Яглом对平稳过程理论所做的一个较系统而又初等的介绍[23]. 对 §24 中提到的平稳广义过程, 可参看本书作者的论文[23] 及И. М. Гельфанд[18]、K. Urbanik[34]、郑绍濂[7][8] 等人的著作. 关于 §31 中所提到的多维平稳过程, 最近有了较多的发展, 可

参看 N. Wiener 与 P.Masani 的著作[35] 及Ю. А. Розанов的著作[32]. 对于本节中提到的齐次随机场的理论, 可参见江泽培的著作[4].

对平稳过程的统计理论有兴趣的同志, 除了可参看本书后记中提到的文献外, 尚可参看 Ulf. Grenandar与M. Rosenblatt 的著作[10]. 我国的数学工作者近几年来在平稳过程 (包括多维平稳过程及齐次随机杨) 的回归系数的估计方面也进行了一定的工作, 这可参看王寿仁[3]、江泽培[5]、郑绍濂与陶宗英[9] 等人的著作.

第 4 章① §43, 可参看Е. Б. Дынкин与А. А. Юшкевич[12], А. А. Юшкевич[33] 及 R. M. Blumenthal[10] 等人的著作. §45, 可参看Е. Б. Дынкин的著作[13]. §48, 可参看 S. Karlin 与 J. McGregor[26] 和王梓坤[1][2] 等人的著作.

第 5 章 §51 可参看 W. Feller 的著作[15], §52 可参看Е. Б. Дынкин的著作[14], §53, 可参看Е. Б. Дынкин[13] 和 W. Feller[16] 的著作. §57 可参考 D. A. Darling 与 A. J. Siegert[11] 和Р. З. Хасьминский[20] 的著作, §59, 可参考 W. Feller[17] 和Е. Б. Дтнкин[14] 的著作及在此二文中提到的参考文献.

参 考 文 献

[1] 王梓坤：一个生灭过程, 科学记录, 新辑, 8(3), 1959, 266 ~ 268.

[2] 王梓坤 (Ван Цзы-кун): Классификация всех процессов размножения и гибели, Научные Доклады Высшей Щколы, Физ-матем., 4, 1958, 19~25.

[3] 王寿仁：格子点上随机场的回归系数的估计问题, 数学学报, 2(8), 1958, 210~222.

[4] 江泽培：(Цзян Цзз-пей): О линейной экстраполяции непрерывного однородного случайного поля, Теор. Вероятн. и ее Щрим., 1(2), 1957, 60~91.

[5] 江泽培 (Chiaug Tsc-pei): On the estimation of regression coefficients of a continuous parameter time series with a stationary residaul,

① 第 4 章和第 5 章的参考文献是由王梓坤同志提供的.

Теор.Вероятн. и ее Прим., 4(4), 1958, 405~423.

[6] 冯 康：广义函数论, 数学进展, 3(1), 1955, 405~590.

[7] 郑绍濂：正则与奇异的平稳广义随机过程, 科学记录, 新辑, 8(3), 1959, 281~286.

[8] 郑绍濂：多维平稳广义随机过程的谱分析, 复旦大学学报, 2, 1960, 203~210.

[9] 郑绍濂、陶宗英等：具有多维平稳随机扰动的回归系数的估计, 即将发表.

[10] R. M. Blumenthal: An extended Markov property, Trans. Amer. Math. Soc., 1, 1957, 85.

[11] D. A. Darling; A. J. Siegert: Ann. Math. Stat., 24, 1953, 624.

[12] Е. Б. Дынкин; А. А. Юшкевич: Строго Марковские продессы, Теор. Вероятн. и ее Прим., 1(1), 1956, 149~154.

[13] Е. Б. Дынкин: Инфинитезимадьные Операторы Марковских процессов, Теор. Вероятн. и ее Прим., 1(1), 1956, 38~60.

[14] Е. Б. Дынкин: Одномерные непрерывные строго Марковские процессы, Теор. Вероятн. и ее Прнм., 1(4), 1959, 3~54.

[15] W. Feller: On differential operators and boundary conditions, Communication on Pure and Applied Math., 8, 1955, 203~216.

[16] W. Feller: The general diffusion operator and positivity preserving semi-groups in one dimension, Ann. of Math., 3(60), 1954, 417~436.

[17] W. Feller: Diffusion processes in one dimension, Trans. Amer. Math. Sic., 77, 1954, 1~31.

[18] И. М. Гельфанд: Обобщенные случайнле процессы, ДАН СССР, 5(100), 1955, 852~856.

[19] U. Grenander and M. Rosenblatt: Statistical Analysis of Stationary time Series, Wiley, 1957.

[20] Р. З. Хасьминский: Распределение вероятностей для функционалов от траекторин сдучайного процесса диффузионного типа, ДАН СССР, 1(104), 1955, 22~25.

[21] А. Я. Хинчин: Теория корреляции стационапнышl стош-

астическиш процессов, УМН 5, 1938, 42 或 Math. Ann., 109, 1934, 604~615.

[22] K. Itô: Stationary random distributions, Mem. of the Coll. of Sci. Kyoto Univ., Ser. A, 3(28), 1954, 209~223.

[23] A. M. Ярлом: 平稳随机函数导论 (梁之舜译), 数学进展, 1(1), 1956.

[24] K. Karhunen: Über lineare Methoden in der Wahrscheinlichkeitsrechnung, Ann. Acad. Sei. Fennicae, I. Math.-Physica, 37, 1947.

[25] K. Karhunen: Über die Struktur stationärer zufärer Funktionen, Ark. Mat. 1, 1940, 141~160.

[26] S. Karlin; J. McGregor: The classification of birth and death processes, Trans. Amer. Math. Soc., 86, 1957, 366~400.

[27] J. R. Kinney: Continuity properties of sample functions of Markov processes, Trans. Amer. Math., 74, 1953.

[28] A. Kolmogoroff: Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung, Math. Ann., 104, 1931, 415~458.

[29] A. H. Колмогоров: Стационарных последовательности в ьильбертовом пространстве, Вюлл. МГУ 2(6), 1941, 1~40.

[30] Ю. В. Прохоров: Шодимсств случаибых процессов и предельные теоримь терии вероятностей, Теор. Вероятн. и ее Прим., 2(1), 1956, 117~238.

[31] A. B. Скороход: Предельные теоримы для случайных процессов с независимыми приращениями, Теор. Вероятн. и ее Прим., 2(2), 1957, 145~177.

[32] Ю. А. Розанов: Спектральная теория многомерных стационарных цлучайных процессов с дискретным временем, УМН, XШ, 2(80), 1958, 93~142.

[33] A. A. Юшкевич: О строго Марковских процессах, Теор. Вероятн. и ее Прим., 2(2), 1957, 187~217.

[34] K. Urbanik: Generalized stochastic processes, Studia Math., XVI

(1), 1958, 264~334.

[35] N. Wiener; P. Masani: The prediction theory of multivariate stochastic processes, I, Acta Math., 98, 1957; II, Acta Math., 99, 1958, 93~137.